NATIONAL TECHNICAL UNIVERSITY OF ATHENS

Development of data-driven ANN model for the Propeller shaft power prediction & fouling analysis of vessel.

GEORGIOS-AGISILAOS EMMANOUILIDIS

Diploma Thesis

School of Naval Architecture and Marine Engineering

Supervisor: Assistant Professor Nikolaos Themelis

Committee member: Professor Konstantinos Spyrou

Committee member: Professor Ioannis Prousalidis

Acknowledgments

After five years of dedication and studying, I have come to this moment to write this study which serves as the final requirement for closing this journey and getting my engineering diploma from the School of Naval Architecture and Marine Engineering of the National Technical University of Athens. Therefore, I would like to express my gratitude to the people that helped me to achieve this goal.

First and foremost, I would like to thank my supervisor Ass. Prof. Nikolaos Themelis for giving me this opportunity, to proceed with this choice of project and for all the time that he dedicated to helping me and guiding me throughout this project.

Finally, to all my friends, colleagues, and family, I would like to thank you for encouraging me to always be the best version of myself and for assisting me with your unbelievable support in times needed and especially during the completion of this diploma thesis as well as throughout my studies.

Περίληψη

Λόγω των πρόσφατων γεγονότων στον ναυτιλιακό τομέα, η παρακολούθηση της απόδοσης των πλοίων είναι πιο σημαντική από ποτέ. Η σωστή αξιολόγηση της κατάστασης του πλοίου επιτρέπει στις ναυτιλιακές εταιρείες να αξιολογήσουν την υποβάθμιση της απόδοσης και να αποφύγουν απρόβλεπτες δαπάνες. Η απόδοση του πλοίου εξαρτάται από τον χρόνο λόγω της σχέσης του με τη θαλάσσια ρύπανση. Επομένως, τα δεδομένα που καταγράφονται από τα συστήματα παρακολούθησης του πλοίου μπορούν να αξιολογήσουν την κατάσταση της απόδοσής του και να καθορίσουν εάν το πλοίο πρέπει να υποστεί συντηρητικές ενέργειες.

Επιπλέον, η συνδυασμένη χρήση υπάρχοντων καταγεγραμμένων δεδομένων με τεχνικές μηχανικής μάθησης δίνει τη δυνατότητα στους φορείς να παρακολουθούν όχι μόνο την τρέχουσα αποδοτικότητα της απόδοσης ενός πλοίου, αλλά και να προβλέπουν τις διακυμάνσεις της με την πάροδο του χρόνου. Η παρούσα μελέτη περιγράφει τα απαιτούμενα βήματα για την ανάπτυξη ενός προγνωστικού μοντέλου Τεχνητού Νευρωνικού Δικτύου (Artificial Neural Network - ANN) και τα στοιχεία που περιλαμβάνει. Επιπλέον, αξιολογεί τη σημασία κάθε υπερπαραμέτρου σχετιζόμενη με την ακρίβεια του μοντέλου.

Για την ανάπτυξη του μοντέλου ANN, επεξεργάστηκαν δεδομένα υψηλής συχνότητας από ένα Panamax φορτηγό πλοίο για μια περίοδο 16 μηνών, με έμφαση στην επισκευή του προπέλας. Δημιουργήθηκαν δύο ξεχωριστά μοντέλα: ένα χρησιμοποιώντας δεδομένα εκπαίδευσης που καταγράφηκαν κατά την πρώτη εξάμηνη περίοδο και ένα δεύτερο χρησιμοποιώντας τα υπόλοιπα (10 μήνες). Για την προεπεξεργασία των καταγεγραμμένων δεδομένων, εφαρμόστηκαν στατιστικές μέθοδοι μαζί με τους βασικούς νόμους της φυσικής. Η μηχανική των μεταβλητών εισόδου περιλάμβανε την εφαρμογή τεχνικών βαθιάς μάθησης με τη χρήση της βιβλιοθήκης Random Forest της Python.

Για τη βελτιστοποίηση της απόδοσης του μοντέλου, εφαρμόστηκε η μέθοδος δοκιμής και σφάλματος (Trial-Error) για κάθε υπερπαράμετρο. Έτσι, κατασκευάστηκε το πιο αποδοτικό μοντέλο δυνατό, λαμβάνοντας υπόψη τα καταγεγραμμένα δεδομένα, με ποσοστά ακρίβειας 97,5% και 99,3% για τις προβλέψεις σε σύγκριση με τις μετρημένες τιμές πριν και μετά την επισκευή του προπέλα αντίστοιχα. Τέλος, με βάση αυτά τα μοντέλα, η μελέτη αξιολόγησε την επίδραση της περιβαλλοντικής μόλυνσης στην υποβάθμιση της ισχύος του προπέλα, σε σχέση με τις ημέρες που παρήλθαν από την επισκευή του, προκειμένου να αποδειχθεί η χρονική εξάρτηση της θαλάσσιας μόλυνσης.

Abstract

Due to recent events in the shipping industry, vessel performance monitoring has become more important than ever. Correctly evaluating the vessel's condition enables shipping companies to assess performance deterioration which can save them from unpredicted expenses. Ship performance is time-dependent due to its correlation with marine fouling. Hence, data recorded from vessel's monitoring systems can evaluate its state of performance and whether the vessel should go through maintenance actions.

Furthermore, combining existing recorded data with machine learning techniques give operators the opportunity not only to monitor the current performance efficiency of a vessel, but also to project its fluctuation over time. This study outlines the necessary steps needed to develop an Artificial Neural Network predictive model and what it consists of. Additionally, it evaluates the importance of each hyperparameter while relating them to the model's accuracy. Thereby, processing high-frequency data of a Panamax bulk carrier, over a 16-month period, consisting of a propeller repair, led this study to develop one ANN model used to predict the propeller's shaft power for each one of the propeller's conditions. Hence, one model is developed through training data recorded over the first 6-month period while the second one used the remaining data, accounting for 10 months.

For the preprocessing of recorded data, statistical methods along with fundamental principles of physics were applied while for the feature engineering deep learning techniques were implemented using Python's Random Forest libraries. Moreover, in order to tune the model's hyperparameters the Trial-Error procedure was implemented for each one resulting in building the best-performing model possible given the recorded data achieving 97.5% and 99.3% accuracy percentages between predicted and measured values on the models before and after the propeller repair respectively.

Finally, in respect to these models this study assessed the effect of fouling on the vessel's propeller's shaft power deterioration, comparing the time elapsed since the propeller's repair. This analysis aimed to illustrate the time dependency of marine fouling and its impact on the vessel's performance.

Table of Contents

List of Figures

List of Tables

1 Introduction

The history of the shipping industry goes back to ancient times when global trade was based on sailing boats powered by wind. These vessels were used to transport goods to people across the world and played a vital role in the development of global trade and economy. Due to the industrialization and globalization of the world, shipping developed to be the leading industry which provides for almost 90% of global trade. This is mainly because shipping achieves a low energy consumption and, thus, a low cost per unit of carried cargo, making it possible to reach the economies of scale.

1.1 Ship Performance Development

Naturally, as the need for the transportation of goods evolved, so did the need for bigger and faster ships for the shipping companies to be able to comply with the rapidly growing demand. Hence, shipping companies emphasized investing in the development of their vessel’s technology to increase the capacity, speed, and traveling radius. Additionally, a great amount of research was put in the vessel performance monitoring systems. This became very popular since the performance of the ship was closely connected to the vessel’s voyage expenses, such as fuel consumption, etc. Fuel consumption is the main expense of each voyage; thus, it has become a critical issue for shipowners as they try to maximize their profits.

Due to recent events, such as new MARPOL and IMO regulations, fuel has grown into an even bigger percentage of the vessel’s operational expenses. Not only did the MARPOL conventions oblige shipowners to use fuels with less Sulphur content (e.g., LSFO, MGO, MDO, etc.), which are naturally more expensive than traditional heavy fuel oil (HFO) in order to reduce greenhouse gas emissions, but also the fuel prices have increased rapidly due to the embargo of the Russian fuel and other commodities.

1.1.1 Vessel marine fouling

The amount of fuel used must be considered when calculating a vessel's operating costs. Particularly for larger ships that use a lot of fuel during their journeys, fuel expenditures can make up a sizeable part of a vessel's operating costs. The size, speed, and age of the vessel, as well as the kind of engine and fuel utilized, are all factors that have an impact on fuel consumption. The amount of fuel used depends on several additional elements, including the weather, the state of the water, and the weight carried by the vessel.

The cost of operating a vessel can be significantly reduced by reducing fuel usage. One option to do this is to increase the ship's fuel economy, for instance by upgrading the engines or optimizing the design. Using more fuel-efficient operating procedures, such as lowering speed or optimizing the vessel's path to reducing fuel use, is another option to cut back on fuel use. Reducing fuel use can save money while also helping the environment by lowering greenhouse gas emissions and other pollutants. Overall, because it can significantly affect the vessel's operating costs and profitability, fuel consumption is a crucial aspect that vessel operators must take into account while planning and carrying out voyages.

Ship’s consumption, according to (Yusim & Utama, 2017), is heavily related to the ship’s drag forces which are affected by the vessel’s hull fouling - biofouling. Marine fouling is generated by the buildup of micro - and macro-organisms onto the ship’s hull, physically increasing the displacement of the vessel and disturbing the ship’s flow through water. Marine fouling generates surface roughness which increases the drag resistance of a ship moving through water. Studies have shown that intense biofouling growth on a ship's underwater hull can reduce speed by up to 20% and increase fuel consumption by around 50% over the course of a year of operations.

Following (Arndt et al., 2021) research on the generation of biofouling on ships’ hulls is influenced by several different factors, such as:

1. **Design and construction of the vessel:** This is a major factor that influences the hydrodynamic performance of the vessel. Additionally, vessels may be equipped with various technologies such as sea chests, bow thrusters, hull appendages and protrusions, etc.

2. **Operating conditions:** This factor includes operational parameters like the vessel’s operational speed as well as the ratio of time underway compared with time alongside, moored or at anchor.

3. **Trading routes:** The characteristics of seawater vary from ocean to ocean. The temperature, the salinity, and the richness of fouling organisms depend on the route of the vessel and are critical factors for the generation of marine fouling.

4. **Maintenance history:** The maintenance of the vessel is one of the most important, if not the most, factors influencing the vessel’s marine fouling concentration. It includes the age and condition of any anti-fouling coating as well as the vessel’s hull cleaning frequency. There are two primary methods for cleaning a vessel's hull: Dry-Dock cleaning and Underwater cleaning. Although Underwater cleaning is easier, cheaper, and faster, it does not yield the best results. In contrast, Dry-Dock cleaning produces the most effective outcome, but it is a much costlier and time-intensive option. Consequently, it is crucial to discern when it is appropriate to pursue Underwater cleaning versus Dry-Dock cleaning based on the vessel's performance status.

1.1.2 Hull Cleaning

Based on IACS’ publication “Classification Societies-What, Why and How?”, typically, hull cleaning procedures tend to occur in accordance with the intermediate and special surveys of the vessel which are mandatory for the vessel’s seaworthiness. Intermediate inspections include either an out-of-water Drydocking or an Under-Water inspection in Lieu of Drydocking (UNWILD) and are to be completed in the second to third year of each five-year special survey cycle. Additionally, special surveys, or five-year Inspections, are to be completed within five years after the date of build or 5 years after the crediting date of the previous Special Survey – Hull.

Although, regardless of these mandatory cleanings, some operators may issue additional ones according to their vessel’s hull and propeller condition. According to (Adland et al., 2018), vessels that sail in high-temperature waters, tend to generate marine fouling faster than others and may be better to perform an additional underwater cleaning which, depending on the degree of marine fouling, vessel size, vessel segment, operation profile and trading areas, could result to a reduction in the range of 1% to 5% on main engine fuel consumption. Hence, the cost of the underwater cleaning which varies from \$5.000 to \$50.000 could eventually be less than the amount saved on fuel. All these scenarios are not easy to be examined and require close attention to the vessel’s performance status.

1.1.3 Vessel Performance Monitoring Systems (VPM)

To make sure that no unpredicted problems and additional costs come up, marine engineers and vessel operators have introduced new methods for monitoring the vessel’s performance and operational status. These Vessel Performance Monitoring systems (VPM), achieve several significant goals.

Enhancing ship efficiency: VPM enables continuous monitoring of important performance metrics like fuel usage, speed, and engine characteristics. With the help of this data, the vessel's performance can be improved to cut costs and the influence it has on the environment.

Enhancing maintenance: VPM can assist in determining maintenance requirements and avoiding equipment failures, lowering the risk of downtime and related expenses.

Increasing safety: Monitoring a vessel's performance can also serve to increase safety by seeing possible problems like engine overheating or hull damage before they get out of hand.

Ensuring compliance: VPM can be used to track adherence to rules governing emissions, fuel use, and other performance parameters, lowering the possibility of fines and legal repercussions.

Overall, as (Valchev et al., 2022) state, VPM gives vessel owners and operators insightful information about their ship's performance, empowering them to make wise choices and take preventative action to increase effectiveness, safety, and compliance. Hence, it is the shipowner’s best interest to invest in technologies, such as vessel performance monitoring systems that would consecutively make their vessels perform as best as possible and consecutively reduce their fuel consumption. Vessel performance monitoring (VPM) is based on the extensive analysis of the vessel’s gathered data. This analysis can be done by developing different numerical models, which according to the literature, the three main families of numerical models that have been developed and used in the literature are Physical Models (models relying on mechanistic knowledge of the phenomena), Data-Driven Models (models relying on historical data about the phenomena combined with Artificial Intelligence), and Hybrid Models (i.e., a hybridization between Physical and Data-Driven Models).

1.2 Artificial Intelligence & Shipping

Artificial Intelligence and Machine Learning have developed significantly over the last years. The use of AI is growing across various industries, although its capabilities differ. AI has started to affect the logistics sector, by removing monotonous and repetitive jobs, succeeding in having a great potential to speed up and enhance the maritime industry. Specifically, machine learning technology combined with big data analysis have the ability to influence a business' capacity for prediction and improve the effectiveness of its operations. Real-time analytics, better scheduling, automated procedures, and other implementations are some examples.

Combining all the years of research along with the assistance of advanced computing power and wider data availability, prediction models have evolved from basic statistical models to deep learning models, which achieve high accuracy prediction percentage. Classical statistical models, such as linear regression and logistic regression base their predictive abilities on assumptions about data distribution as well as the variables correlation. Through the extensive use and examination of these statistical predictive models, machine learning science developed machine learning prediction models, which in contrast with the statistical models, they relied less on presumptions about the distribution of the data and the relationships between the variables. The purpose of these models, which include decision trees, support vector machines, and random forests, is to find patterns in data and generate predictions based on those patterns. Lastly, deep learning models came along with the development of deep learning algorithms, such as neural networks. These models are used for bigger data sets since they have the capability to acquire knowledge from vast volumes of data and recognize intricate patterns that were once challenging or unfeasible to discern.

The use of machine learning and predictive models have revolutionized many industries along with shipping. Using these models, shipping operations can be optimized. Specifically, according to (Farag & Ölçer, 2020) study, regarding the development of ship performance models based on ANN and regression methods, AI offers the ability to shipping companies to reduce costs in various ways. Firstly, it gives the ability to predict when a vessel might need maintenance, by analyzing existing data. Additionally, route optimization is an important aspect of machine learning use in the shipping industry, since shipping companies can decrease fuel consumption, cut costs, and enhance delivery times by scrutinizing data on factors that impact shipping durations such as weather conditions, traffic, and other variables, and subsequently determining the most efficient routes.

1.3 Literature Review

Kim et al., (2021) study, present certain models that can predict fuel consumption. Data from a container ship were used to develop these models, and for their implementation, Artificial Neural Networks (ANN) and Multiple Linear Regression (MLR) were used. Additionally, for the development of this specific model, unlike other studies that use fuel consumption per hour as the dependent variable, fuel consumption per distance traveled was used. This way, it takes into account the conditions under which the ship is sailing, such as the loading condition and the weather. Then, for the selection of the dependent variables to be used for the model development, the "Domain Knowledge" and "Statistical method based on Lasso Regularization" methods were examined. Thus, for the evaluation of the methods, results were obtained for experiments with all possible combinations of variable selection and model training methods, and the mean absolute error of each combination was calculated. Applying the above, it was found that the optimal combinations were the use of Artificial Neural Networks in combination with either the "Domain Knowledge" method or the "Lasso Regularization" method. Therefore, the parameters selected with the "Domain Knowledge" method are SOG, RWS, RWD, DFT, TRM, and DBS, while with "Lasso Regularization" they are RPM, SOG, STW, RWS, RUD, and DBS. Finally, a sensitivity analysis of the model was performed regarding the ship's draft. For this analysis, the combination of ANN - Domain Knowledge was used, which has the draft of the ship as a variable with which the optimal draft of the ship was calculated concerning the ship's consumption.

Furthermore, Lang et al., (2022) made a comparison of various artificial intelligence and statistical methods used for predicting the necessary propulsion power of a ship to achieve operational speed. Specifically, the algorithms compared include XGBoost, ANN, Support Vector Regression, and statistical methods such as Linear and Polynomial Regression and Generalized Additive Model. The parameters used as input variables in these models for comparison are ship speed through water, mean draft, trim, heading, significant wave height, mean wave period, mean wave heading, and wind speed. These variables are used as input variables in computational models that generate the propulsion power as a result. The root means square error and the complexity degree of the method are used to compare these methods. After collecting and processing data from a tanker and a RoRo vessel, it was initially found that artificial intelligence algorithms offer greater reliability and accuracy compared to statistical methods. However, the text notes that statistical methods are superior in the time required for "training". Finally, it turns out that the optimal algorithm regarding the above parameters is XGBoost, for which a sensitivity analysis of the model is also performed regarding the time intervals in which we group our data, and it was concluded that 30 minutes is the best choice.

In Laurie et al., (2021) paper various models of artificial intelligence are being examined, aimed at predicting propulsive power and analyzing pollution in relation to efficiency reduction. Specifically, the following models are being compared: Multiple Linear Regression, Decision Tree (AdaBoost), K-Nearest Neighbors, ANN, and Random Forest. To develop this specific model, data was collected from a container transport ship operating between Europe and South America, an area where the waters have higher temperatures and therefore pollution development is more intense. Additionally, the ETR (Extra Trees Regression) model was used to select the optimal combination of variables, leading to the following variables: behavior, draft, Froude number (Fn), wind speed ratio, significant wave height, water temperature, and days since the last tank cleaning (DSC). Root Mean Square Percentage (RMSPE) and Mean Absolute Percentage Error (MAPE) values were calculated to compare the computational models. Using these values, it was found that the Random Forest model performs the best, although the KNN and ANN models were equally good. Therefore, it appears that non-linear relationships between inputs and outputs work better than linear relationships.

Additionally, a study by Uyanık et al., (2020) is based on determining the fuel consumption of a container using various artificial intelligence algorithms. Specifically, different predictive models such as Multiple Linear Regression, Ridge and LASSO Regression, Support Vector Regression, Tree-Based Algorithms, and Boosting Algorithms were examined using noon reports and data from engine log books. For the validation of the predictive models, the K-fold cross-validation method was used. For the analysis of the relationship between each variable correlation analysis is used, calculating the correlation of all possible variable pairs. Additionally, for the evaluation of each method regarding its accuracy, a root mean square error analysis was used, reporting that it is more advantageous for larger datasets. After gathering all the information, it was concluded that the best-performing, and most accurate predictive model is considered to be the Gradient Boosting Regression.

Karagiannidis, (2019), extensively examined the impact of data pre-processing while creating data-driven models for ship propulsion. In order to train models that forecast the required shaft power or main engine fuel consumption for a container ship sailing under random conditions, he employed a sizable, autonomously acquired data set with a high sampling frequency. Two strategies were proposed with the aim of highlighting the statistical evaluation and preparation of the data. Additionally, state-of-the-art training and optimization methodologies for Feed-Forward Neural Networks (FNNs) were applied. His findings suggest that the accuracy of the model can be significantly increased by a diligent filtering and preparation stage. In addition to that, Karagiannidis & Themelis, (2021), conducted an article regarding the effect of data pre-processing on the prediction of ship fuel consumption and speed loss and ultimately concluded that with the appropriate pre-processing and filtering of data, it is promising to achieve an increased model accuracy.

Lastly, Anastasiou, (2022) through his diploma thesis, made a comprehensive analysis on the development of ship performance models using Artificial Neural Networks and operational data. Specifically, he developed an artificial neural network as well as a multiple linear regression model which were used to predict the fuel oil consumption of a vessel. In his study he examined multiple error detection methods, as well as regularization techniques to overcome overfitting. He mentions that due to the lack of use of engine-related parameters in the designed model, it was not easy to achieve a low error index. Additionally, he highlighted the fact that deep learning models are very likely to come across overfitting problems, due to the large amount of data that are being processed. Lastly, he concluded that a deep neural network using proper parameters, can outperform a multiple linear regression model in many cases and specifically in the prediction of various ship propulsion characteristics.

1.4 Thesis Objective & Structure

As mentioned in this study's introduction, vessel's main engine shaft power monitoring has been of major importance for shipping companies not only for the ship's performance but also for the minimalization of operational expenses and ship's emissions. Its accurate forecasting has been one of the most popular topics in marine engineering and machine learning since an accurate Artificial Neural Network prediction model could be a huge asset for a shipping company since it could eventually cut down their expenses significantly while organizing their vessel's maintenance actions in such a way that the vessel will never end up in a situation where it has broken down entirely. Hence, the objective of this study revolves around developing an Artificial Neural Network model that can effectively forecast a vessel's Shaft Power with high accuracy across different loading and weather conditions. Firstly, to do that it was required to accumulate data from a vessel for a specific period since these models are mainly data-driven (Chapter 2. Data acquisition). Secondly, to create a model capable of producing accurate predictions it is essential to process the given data in order to remain with a filtered Dataset (Chapter 3. Data pre-processing). In the next chapter, Chapter 4. Feature Engineering, the pre-processed dataset's features will be examined while also there will new ones generated. This step is really crucial for the predictive model architecture since the remaining features will serve as the model's input variables. Subsequently, after creating the input dataset along with its final input variables, in Chapter 5. Artificial Neural Networks, the tuning of the ANN model's hyperparameters will be examined and the most suitable model will be created. Finally, after evaluating and selecting the most accurate model, the fouling state of the vessel will be assessed, and the consequences that it has on the vessel's performance efficiency as well as its Operational Expenses. Hence, this study will take a deeper look into the optimization of the hyperparameters tuning of a Neural Network as well as the Features Engineering by using programming optimization tools, such as Random Forest Regression Classifier, which is provided by Python's Libraries. Additionally, it will introduce a method to evaluate and assess the vessel's fouling status along with the additional fuel expenses that are associated with it by correlating the additional Propeller Shaft Power to the Days that have passed by since a specific time. This will indicate a positive correlation between the Days Since the last Repair to the Propeller's Shaft Power.

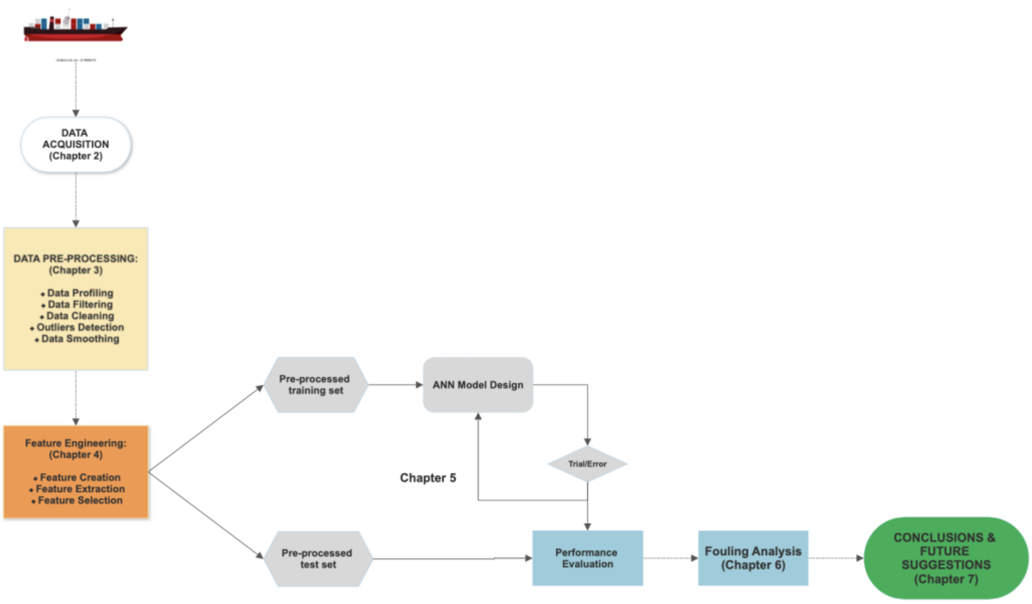


Figure 1: Thesis Structure Flow Graph

2 Data acquisition

For the development of the prediction model, data were collected from the operation of a Bulk Carrier Vessel (MV Kastor) during the period of January-2021 to April-2022. For the collection of the data, a continuous monitoring system was used with a frequency of 1 minute. These data describe several parameters and were captured through various sensors. The total number of data points was 698400. Additionally, during that period the vessel has gone through propeller maintenance which is expected to affect the vessel's performance efficiency indicators.

Vessel's Main Characteristics	
Type	Bulk Carrier
Length (BP)	225.5 [m]
Length (OA)	229 [m]
Beam	32.3 [m]
Draft	14.45 [m]
Depth	20.05 [m]
Deadweight	81600 [tn]
Engine	MAN B&W 6S60ME-C8
	9930 kW / 90 RPM
Service Speed	14 [knots]

Figure 2: Main characteristics of reference vessel.

Data acquisition plays a vital role in vessel monitoring and is essential for creating both statistical and AI predictive models. As (Skamagkas, 2022) states, there are two main methods for the collection of operational data:

<u>*Noon reports:*</u>

The crew of a ship normally submits noon reports each day, informing the shipping company of the vessel's location, speed, heading, and other operational information. All these data, collected from noon reports, can be gathered in a vessel monitoring system in several ways:

Electronic submission: The crew submits the report to the monitoring center electronically using a computer or mobile device. The report will then be automatically processed and saved in the system.

Fax or email: The crew sends a fax or email to the monitoring center or the company responsible for the vessel's operation, which can then manually enter the data into the system.

Manual entry: The data can be manually entered into the system by the monitoring center, either from a paper copy of the report or from an electronic copy that was obtained from the crew.

<u>*Continuous monitoring systems*</u>

These systems collect data with a higher frequency than the noon reports, such as a few seconds. For data collection, these monitoring systems use several sensors and communication systems installed on the vessel.

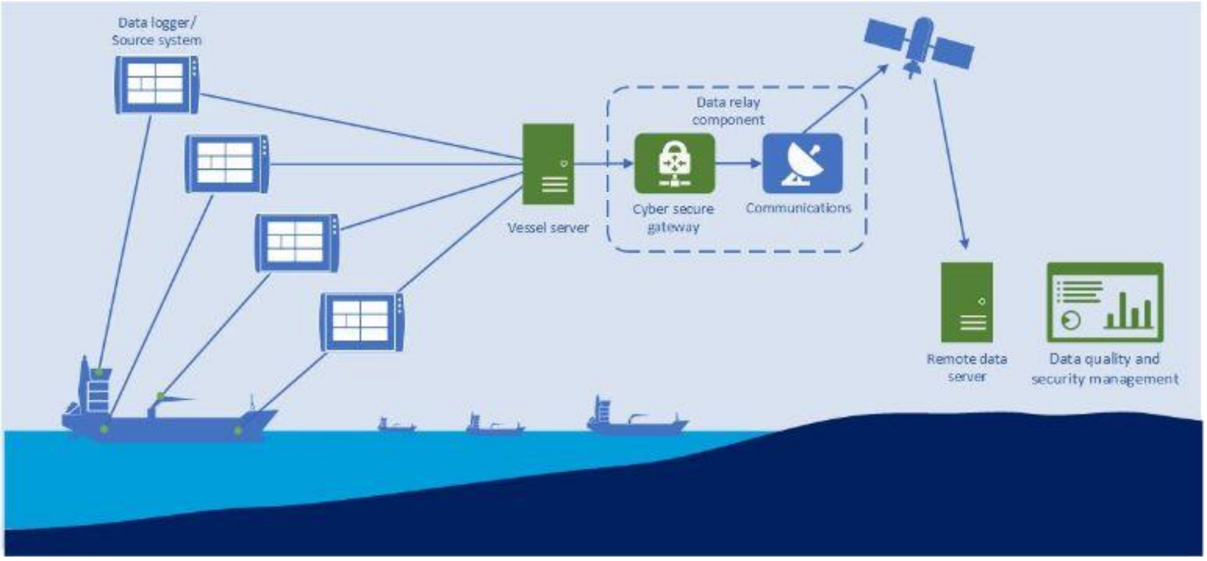


Figure 3: Continuous data monitoring for vessel performance.

(*Digitalization in the Maritime Industry - DNV*, n.d.)

Regardless of the method of collection, both methods can provide us with valuable information about the vessel’s operational status. Although, due to the higher frequency that continuous monitoring systems have, they provide us with a more detailed and more accurate view of the vessel’s operations compared to noon reports.

Parameter	Sensor	Units
Speed Over Ground (SOG)	GPS	knots
Speed Through Water (STW)	Speed Log	knots
Propeller Shaft Power (PSP)		kW
ME Revolutions per Minute	Shaft Torque Meter	RPM
ME Loading percent		%
Fuel Index Position	Mass Flow Meter	N/A
ME Fuel Oil Consumption		mt/day
Vessel Heading	GPS	deg
Relative Wind Direction	Anemometer	deg
Relative Wind Speed		m/sec
DTN AIR TEMPERATURE 10M ACTUAL	Thermometer	°C
DTN AIR PRESSURE MEAN SEA LEVEL ACTUAL		mbar
Draft	Pressure Sensor	m
Vessel Trim		m
DTN SEA TEMPERATURE 0M ACTUAL	Thermometer	°C
Fuel Oil Temperature (ME return)		°C
Fuel Oil Temperature (ME supply)		°C
Shaft Torque	Shaft Torque Meter	kNm
Shaft Thrust		kN
M/E Shaft Revolutions		RPM
Significant Wave Height		m
Mean Wave Direction		deg
Water Depth Relative to the Transducer	Echo sounder	m
Ballast-Condition		N/A
Cargo Carried		tn
Fuel-LCV		kJ/kg

Table 1: Parameters captured through various sensors.

Speed Logs

Speed logs are used to calculate the vessel's speed through water. To that extent, the following sensors are mostly used:

Doppler Logs

Doppler speed logs utilize the Doppler phenomenon to measure speed by detecting the shift in wavelengths of moving objects relative to the observer. This shift is converted into speed by emitting high-power acoustic energy into the water and receiving the echo reflected from the seafloor. The Doppler shift from the returning echo is used to determine the speed of the water passing the sensor, as well as the distance traveled and depth of the water. The sensor is placed on the vessel's longitudinal axis, about 1/3 of the length forward, with the boundary layer of water (whose speed is measured) typically located 2-7m below it. Although, doppler speed logs can malfunction through shallow waters due to the water's acceleration along with the change of direction of the vessel.

Electromagnetic Logs

The electromagnetic log operates by creating an electromagnetic field in the surrounding water through the generation of a small alternating current in a transducer. As the vessel travels through the water, a voltage corresponding to the speed is produced perpendicular to the direction of travel. This voltage is detected by probes and sent to the main electronic unit where it is amplified and processed digitally before being transmitted to the speed and distance displays.

Acoustic correlation Logs

A less common alternative to the aforementioned speed logs. Acoustic correlation logs are based on the correlation of the reflected pulses (sound-energy) in the water at a given distance. The time delay of two similar pulses is measured, and the speed of the ship is calculated.

Echo sounder

An echo sounder, commonly referred to as a depth sounder or sonar, uses sound waves to measure the depth of the water beneath a vessel. It operates by sending out a sound wave that passes through the water and bounces against the seafloor. The length of time it takes for a sound wave to reach the ocean floor and return is measured, and the depth of the water is determined using this length of time. The make-up of the seafloor and the presence of submerged items are two additional pieces of information that contemporary echo sounders may offer.

Shaft Torque Meters

As their name suggests, Shaft torque meters are used to measure the shaft power of the vessel's engine. The two most common ways are the following:

Shaft rotation angle

To measure the torque, the shaft's rotation angle is used, which is calculated through the angular difference between two rings that are placed at a certain distance on the axis. Thus, the angle of rotation is obtained.

Strain Gauges

Strain gauges measure an object's deformation or strain at a 45-degree angle. They are made up of a thin, flexible wire or film attached on a surface that adjusts its resistance in response to pressure or strain. The amount of deformation is inversely correlated with the change in resistance, which can be measured and used to determine the object's strain.

Shaft Torque Meters

As their name suggests, Shaft torque meters are used to measure the shaft power of the vessel's engine. The two most common ways are the following:

Shaft rotation angle

To measure the torque, the shaft's rotation angle is used, which is calculated through the angular difference between two rings that are placed at a certain distance on the axis. Thus, the angle of rotation is obtained.

Strain Gauges

Strain gauges measure an object's deformation or strain at a 45-degree angle. They are made up of a thin, flexible wire or film attached on a surface that adjusts its resistance in response to pressure or strain. The amount of deformation is inversely correlated with the change in resistance, which can be measured and used to determine the object's strain.

GPS (Global Positioning System)

GPS is a system used to track the position of an object. Hence, it retrieves information about the position of the ship in global coordinates (longitude, latitude), and therefore, by numerically extracting the position of the ship, the calculation of the ship's speed over ground (SOG) is achieved. In order for the GPS to work properly it requires to be continuously connected to a satellite system that will transport the ship's location. Specifically, The GPS system consists of more than 30 orbiting navigational satellites. Because they are continually sending out signals, we know where they are. The vessel's GPS receiver watches for these signals. The receiver can determine the vessel's location after calculating its distance from four or more GPS satellites. The accuracy is great but can be affected by currents.

Pressure Sensor

Pressure sensors are installed in vessels for the calculation of their trim and drafts. The draft is calculated by measuring the hydrostatic pressure at the surface of the vessel's hull using sensors placed on the external surface. The draft is then calculated at the location where the sensors are installed. To account for the impact of dynamic changes on drafts, such as the effect of waves, a sensor or measurement of the vessel's movements, such as an inertial measuring unit sensor, is used. Additionally, the trim can be calculated by measuring the draft at two different longitudinal positions of the ship.

Mass Flow Meters

Coriolis Mass Flow Meters

To measure the fuel consumption of an engine, mass flow meters are used. Specifically, Coriolis mass flow meters are known to provide the most accurate results due to the fact that they don’t rely on fuel density estimations since they measure the mass flow directly. They operate on the idea of Coriolis acceleration, which happens whenever a fluid is rotated. A U-shaped tube that vibrates at a resonant frequency makes up the meter. The Coriolis force that the fluid experiences as it passes through the tube causes it to bend and twist. The fluid's mass flow rate directly relates to the degree of deflection. Additionally, the mass flow rate is determined by measuring the time delay between the signals produced by sensors at the tube's input and outlet ends, which detect deflection. The mass flow rate and fluid density can both be determined using Coriolis mass flow meters.

Anemometer

The wind anemometer is a tool that shows the wind's relative direction and speed in relation to the ship's orientation. It is made up of a vane and a helicoid propeller that measure the direction and speed of the wind, respectively. The helicoid propeller's rotating speed and the vane's angular displacement both aid in estimating the relative direction and speed of the wind.

3 Data pre-processing

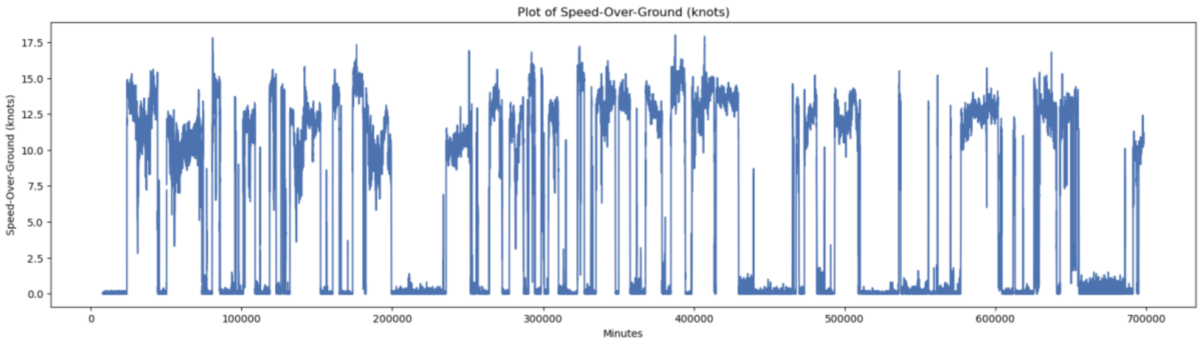
This chapter emphasizes on the preparation of the raw data to prepare it for the main data processing procedure. Data pre-processing is one of the most important steps for the creation of a machine learning model. Real-world data is unorganized and frequently produced, processed, and saved by a range of people, business operations, and software programs. Because of this, a data set can be incomplete, have manual input errors, have duplicate data, or use several names to refer to the same object. In the data that they use for their line of work, humans can frequently spot and fix these issues, but data used to train machine learning or deep learning algorithms needs to be automatically pre-processed. For that reason, the vessel's raw data were processed according to the following steps:

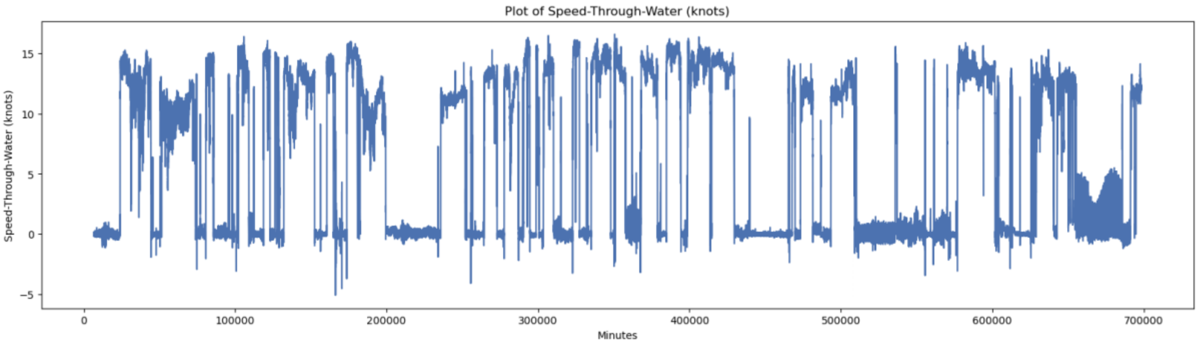
1) **Data profiling**: Examining, evaluating, and reviewing data in order to compile statistics regarding its quality. Hence, it is important to visualize the data by creating plots and histograms which will give us a better understanding of our collected data.

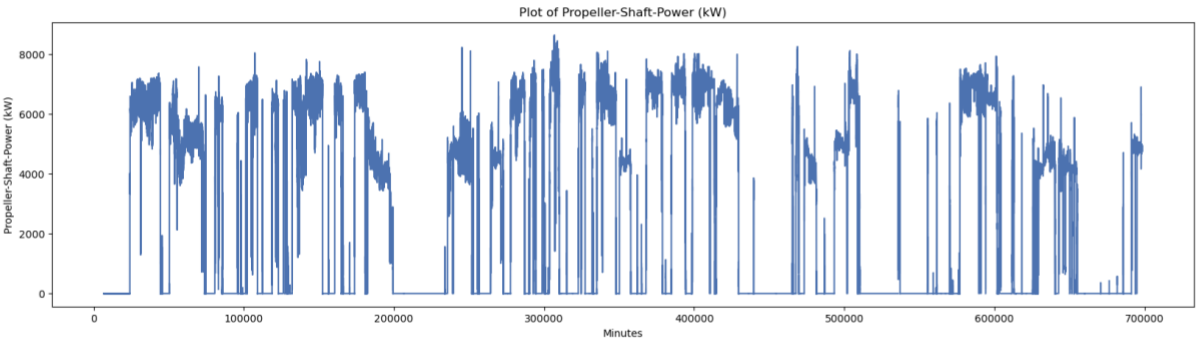
2) **Data filtering:** The purpose here is to determine the easiest solution to remedy quality issues, such as deleting bad data, filling in missing data or otherwise ensuring the raw data is adequate for feature engineering.

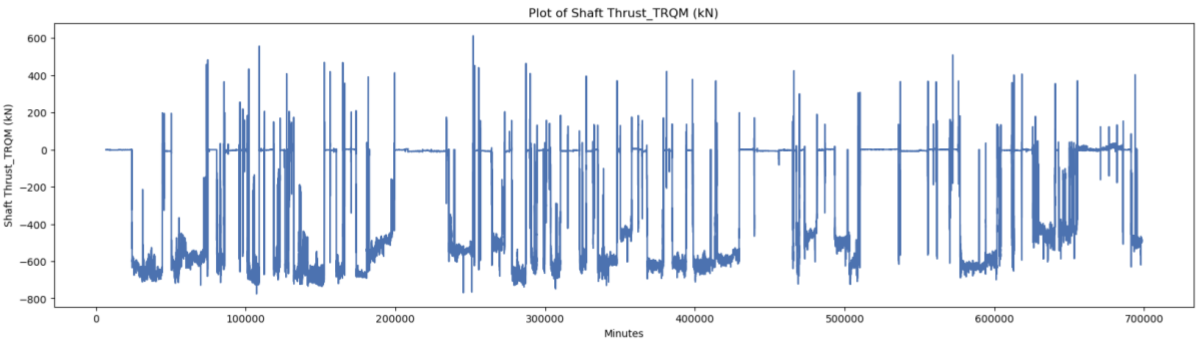
3.1 Data Profiling

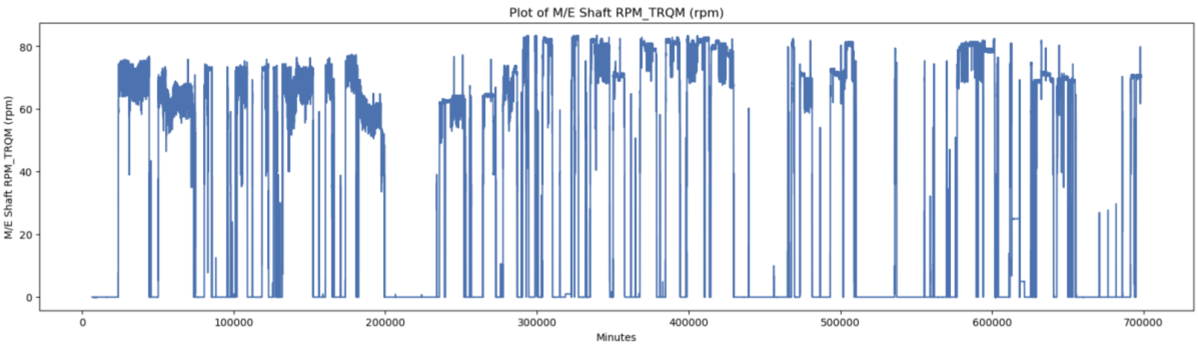
All the collected data from the vessel are time-dependent, hence it is important to have each variable plotted over time to check that they were recorded without any flaws and to detect any sensor failures that may arise. However, given the fact that inside the recorded data there are 29 parameters stored over a period of 22 months, we chose to plot only the most important ones. Ship performance is influenced by three main categories of parameters regarding the vessel's operation, the vessel's loading condition, and the environmental conditions. Hence, in order to demonstrate all of them, representatives for each one were selected. Therefore, Longitudinal Speed Through Water (STW) with ME Fuel-Oil Consumption, Mean Draft with Vessel's Trim, and Wind Speed with Significant Wave Height, were chosen to represent the operational, the loading as well as the environmental parameters accordingly and are presented in Figure 4, Figure 5, and Figure 6 respectively.

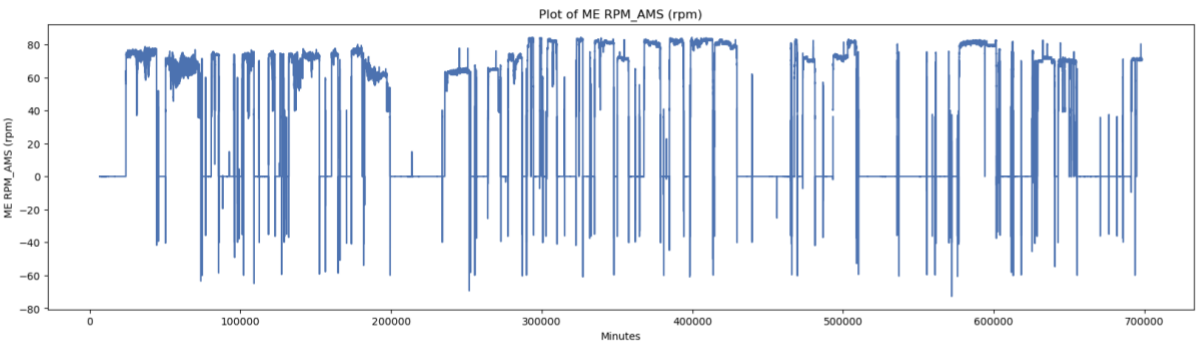



Plot of Speed-Through-Water (knots)
Speed-Through-Water (knots)
15
10
5
0
−5
0
100000
200000
300000
400000
500000
600000
700000
Minutes


Plot of Propeller-Shaft-Power (kW)
Propeller-Shaft-Power (kW)
8000
6000
4000
2000
0
0
100000
200000
300000
400000
500000
600000
700000
Minutes


Plot of Shaft Thrust_TRQM (kN)
Shaft Thrust_TRQM (kN)
600
400
200
0
−200
−400
−600
−800
0
100000
200000
300000
400000
500000
600000
700000
Minutes


Plot of M/E Shaft RPM_TRQM (rpm)
M/E Shaft RPM_TRQM (rpm)
80
60
40
20
0
0
100000
200000
300000
400000
500000
600000
700000
Minutes


Plot of ME RPM_AMS (rpm)
ME RPM_AMS (rpm)
80
60
40
20
0
−20
−40
−60
−80
0
100000
200000
300000
400000
500000
600000
700000
Minutes

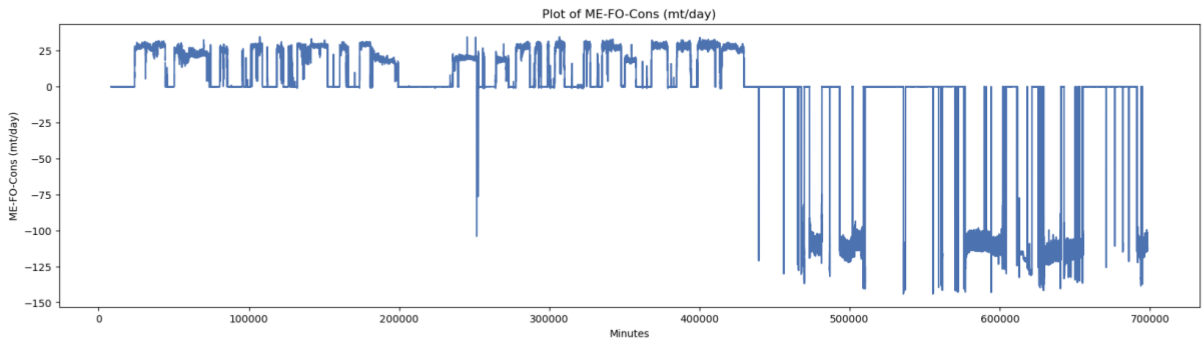


Figure 4: Basic operational parameters over a 22-month period.

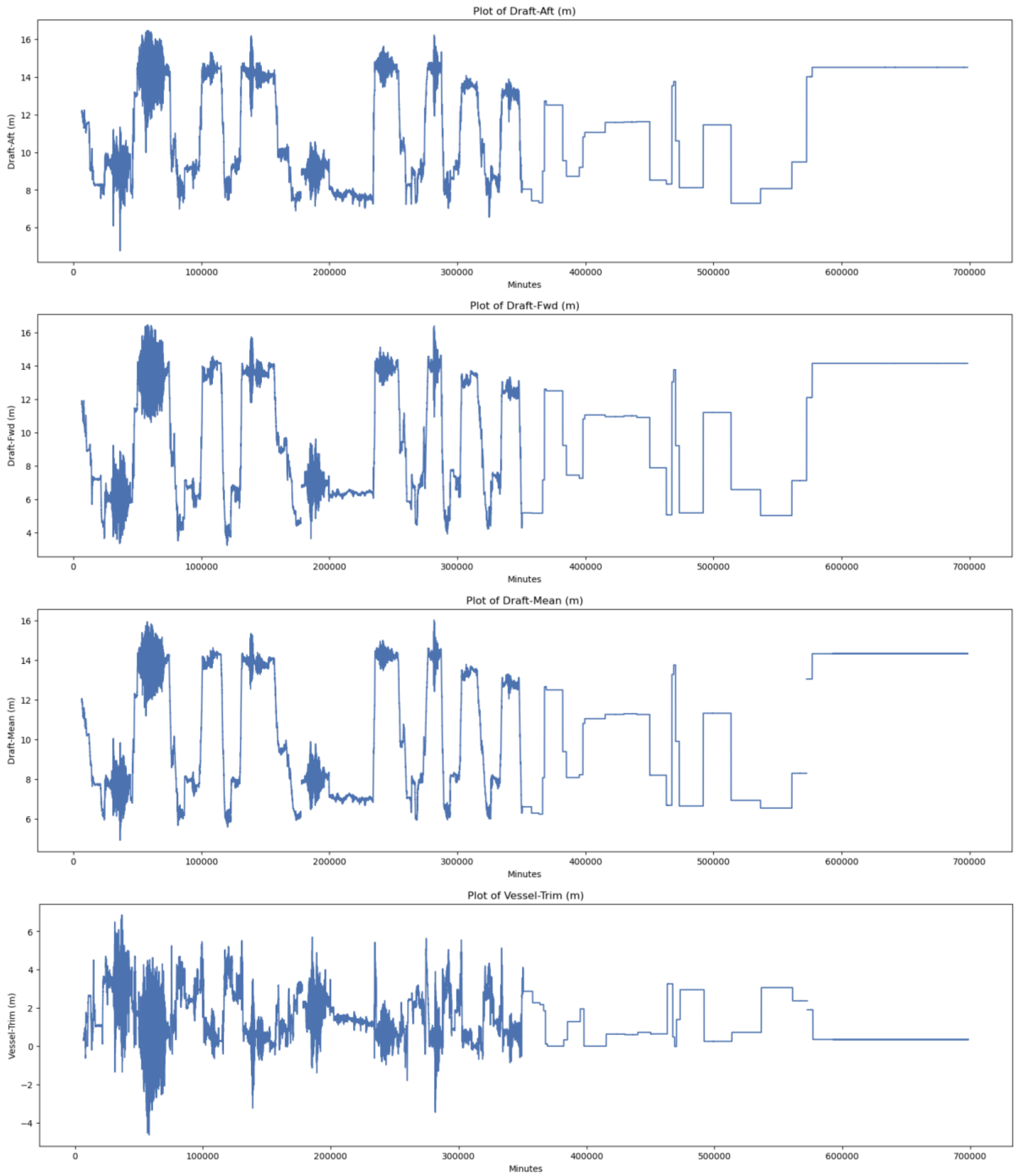


Figure 5: Basic loading parameters (Drafts, Trim) over a 22-month period.

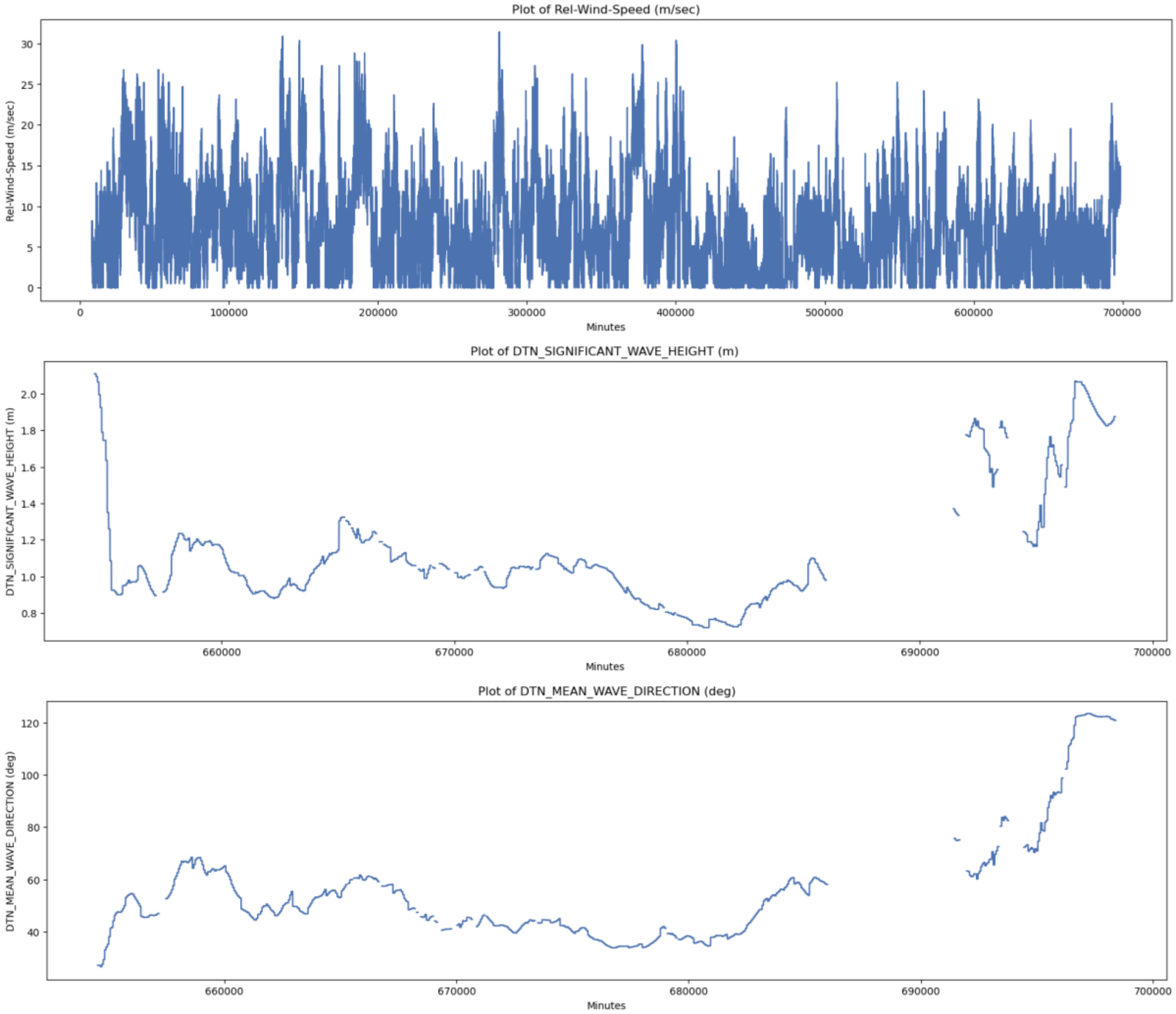


Figure 6: Basic environmental parameters (Relative Wind Speed, Significant Wave Height) over a 22-month period.

From the above-plotted parameters, presented in Figure 4, Figure 5, and Figure 6 the following can be observed:

- The Mass Flow meter sensor malfunctioned through the recorded period, as the collected data showed negative Fuel Oil consumption which is not right. Hence, the ME Fuel-Oil consumption parameter cannot be used for the development of the predictive model.

- In the range of 350000 and 700000 minutes the vessel's pressure sensor did not work properly, since the data that were recorded are not continuous. Therefore, the parameters of the mean draft and the trim of the vessel cannot be used through the entire recorded period.

- It is necessary, through the data visualization, to evaluate the existence of noise and outliers as well as the necessity for filtering or smoothing the data. Since this study aims to develop a predictive model for the vessel's shaft power, it is important to clear useless data that could eventually sabotage our model.

Another important step for the pre-processing of data is the identification of missing values and illogical measurements. The following table (Table 2) shows the missing values for each parameter in the entire data set.

Parameter	Missing values
TIME	0
Speed-Over-Ground (knots)	10056
Speed-Through-Water (knots)	8712
Propeller-Shaft-Power (kW)	10476
ME RPM_AMS (rpm)	8245
ME-Loading-percent (%)	14011
Fuel-Index-Position (n/a)	7597
ME-FO-Cons (mt/day)	14737
Vessel-Heading (deg)	10064
Rel-Wind-Direction (deg)	10546
Rel-Wind-Speed (m/sec)	10099
DTN_AIR_TEMPERATURE_10M_ACTUAL (oC)	656569
DTN_AIR_PRESSURE_MEAN_SEA_LEVEL_ACTUAL (mbar)	656569
Draft-Aft (m)	8841
Draft-Fwd (m)	8891
Draft-Mean (m)	9714
Vessel-Trim (m)	9714
DTN_SEA_TEMPERATURE_0M_ACTUAL (oC)	663293
Fuel Oil Temperature (ME return)_TRQM (C)	7228
Fuel Oil Temperature (ME supply)_TRQM (C)	7188
Shaft Torque_TRQM (kNm)	7169
Shaft Thrust_TRQM (kN)	7170
M/E Shaft RPM_TRQM (rpm)	7167
DTN_SIGNIFICANT_WAVE_HEIGHT (m)	663001
DTN_MEAN_WAVE_DIRECTION (deg)	663001
Water Depth Relative to the Transducer_BRG_ECHO (m)	124335
Ballast-Condition (-)	11154
Cargo-Carried (tn)	8458
Fuel-LCV (kJ/kg)	14632

Table 2: Missing values of the raw data

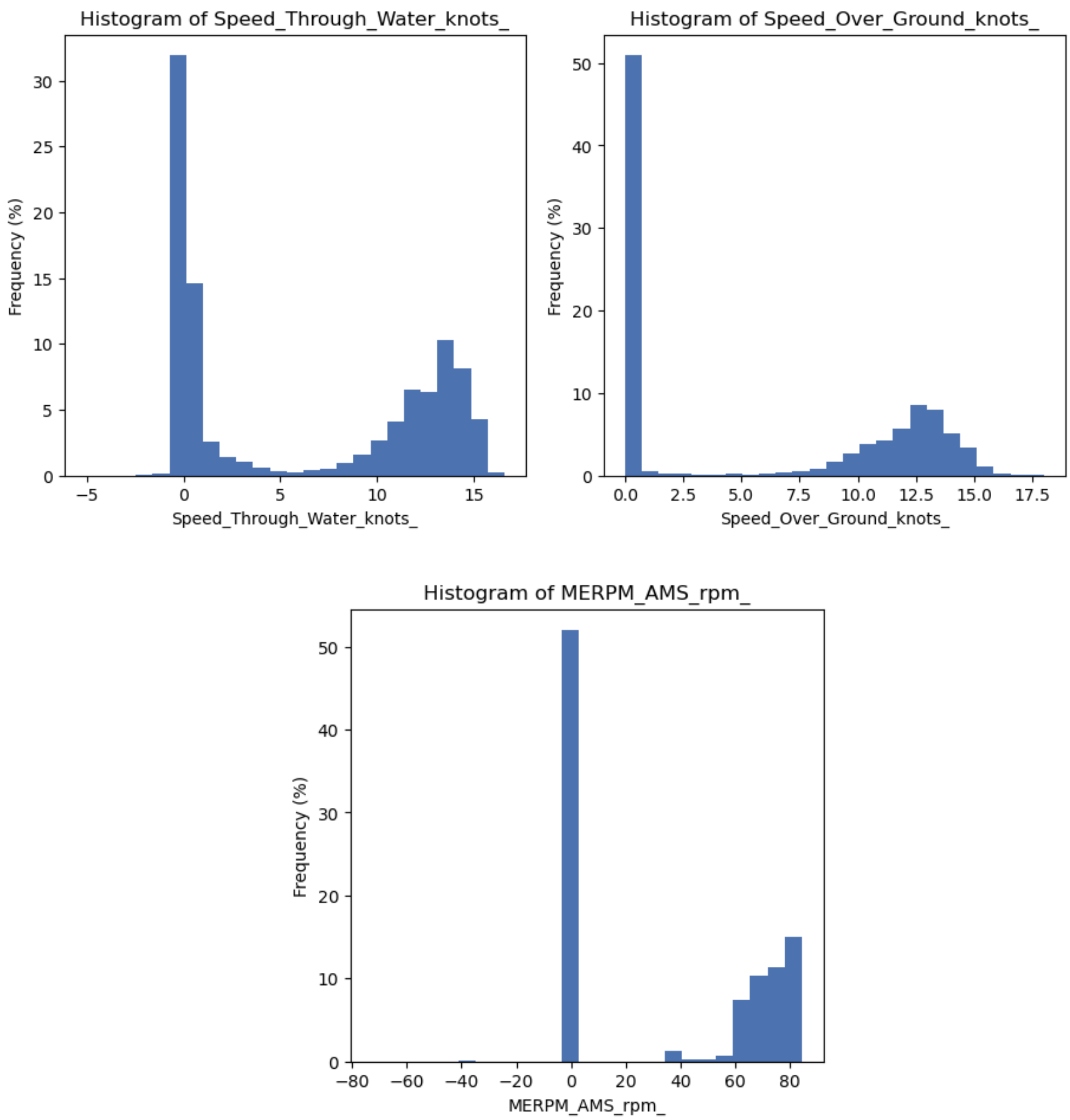
From Table 2: Missing values of the raw data, it can easily be observed that some parameters have too many missing values and therefore it would be wise not to use them. Hence, the following parameters will be deleted from the data set:

1) DTN AIR TEMPERATURE 10M ACTUAL (ºC)
2) DTN AIR PRESSURE MEAN SEA LEVEL ACTUAL (mbar)
3) DTN SEA TEMPERATURE 0M ACTUAL (ºC)
4) DTN SIGNIFICANT WAVE HEIGHT (m)
5) DTN MEAN WAVE DIRECTION (deg)

Apart from the above parameters, the ME Fuel Oil consumption parameter was deleted since we saw that the mass flow meter sensor malfunctioned. After this procedure, we are now left with the parameters that will help us build our model. Additionally, the Water Depth parameter should not be included in the model creation since it is used only in shallow waters.

With a better understanding of the dataset, we can now identify the relationships between its parameters and choose the appropriate processing steps for each. Due to our understanding of the mechanics behind the problem, we already know the relationship, or at the very least the overall trend, for some of them (such as PSP-RPM), which will be useful in reducing outliers.

In order to evaluate each parameter and get a clearer view of their variance, the following histograms were created. These histograms show the relative frequency of different values of each parameter.



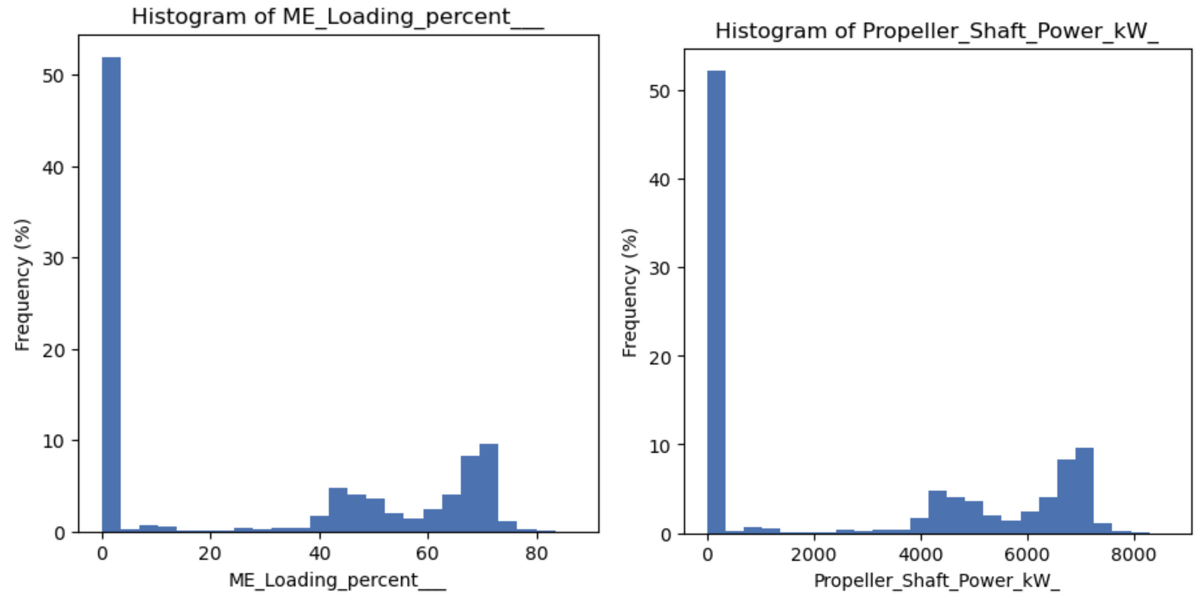
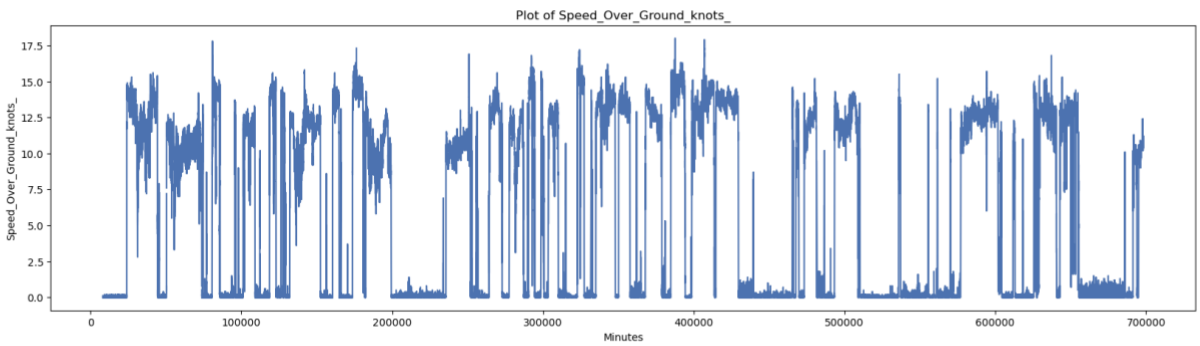
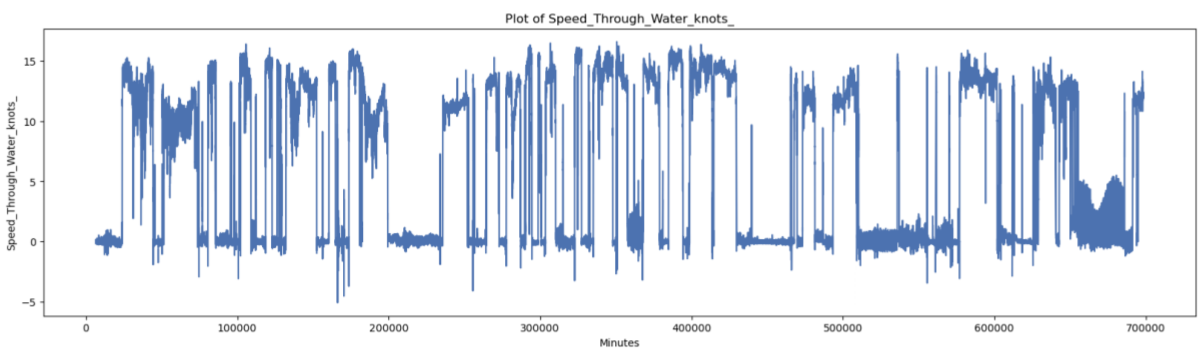


Figure 7: Histograms of relative frequency of raw data parameters

It is noticeable that all parameters have more null values than expected. Also, due to the amount of collected data through the 22-month period, it is difficult to make assumptions for the parameter relationships. Therefore, it would be better if the filtering of the data was initiated at first.

Finally, due to sensor malfunctions, it was deemed necessary to use the data from the noon reports as well. Hence, a new dataset was created which contained both data from noon reports and telemetry data. The merged dataset can be obtained in the following plots:





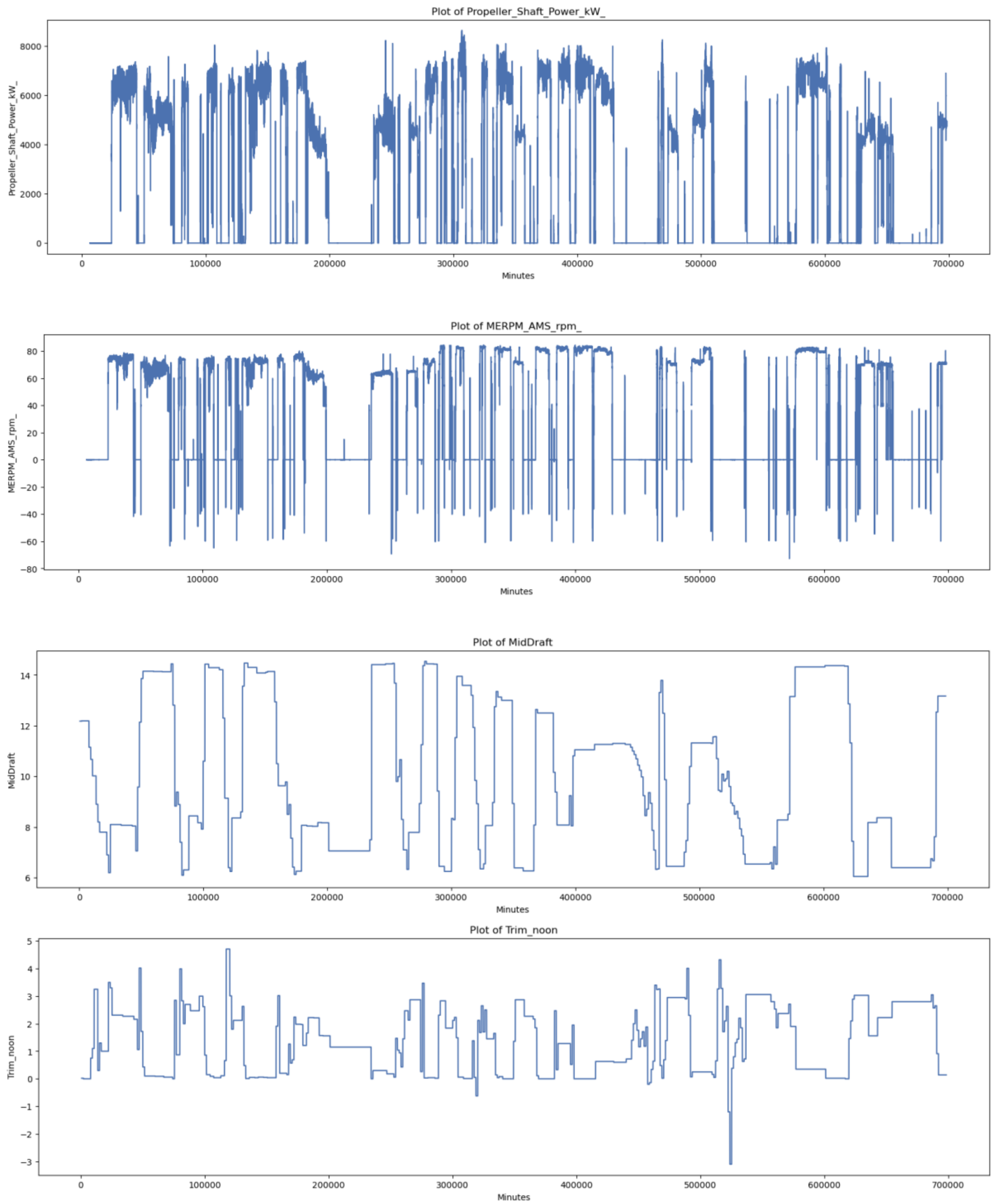
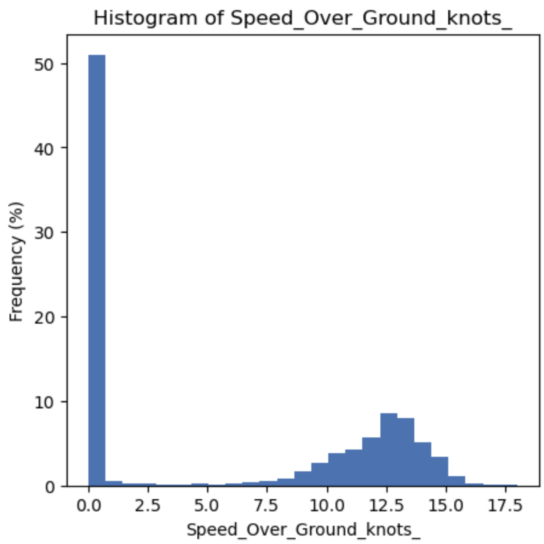
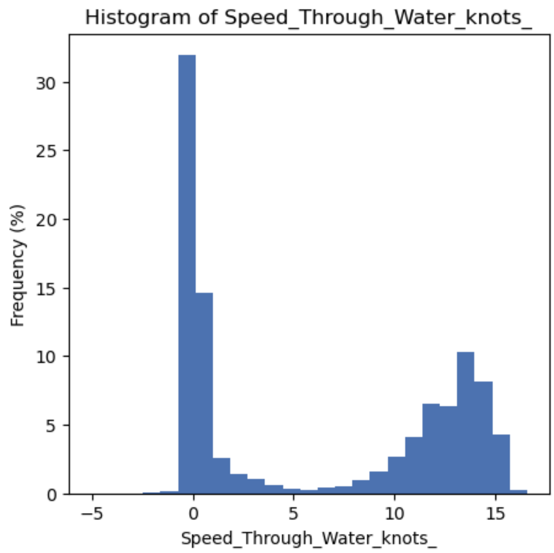
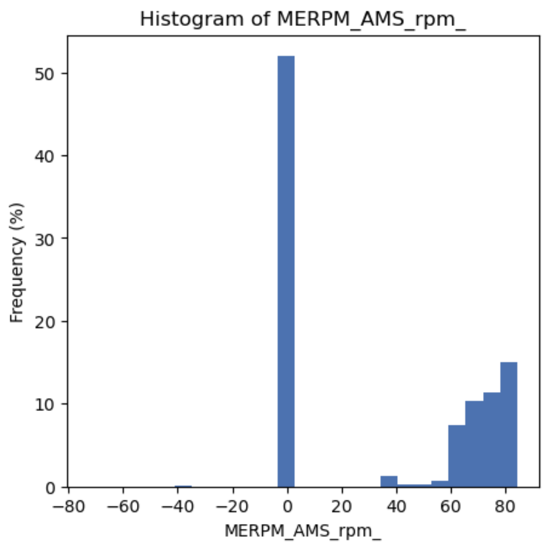
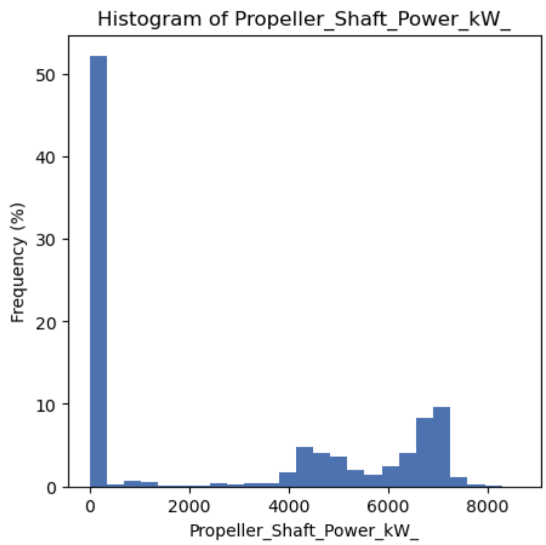


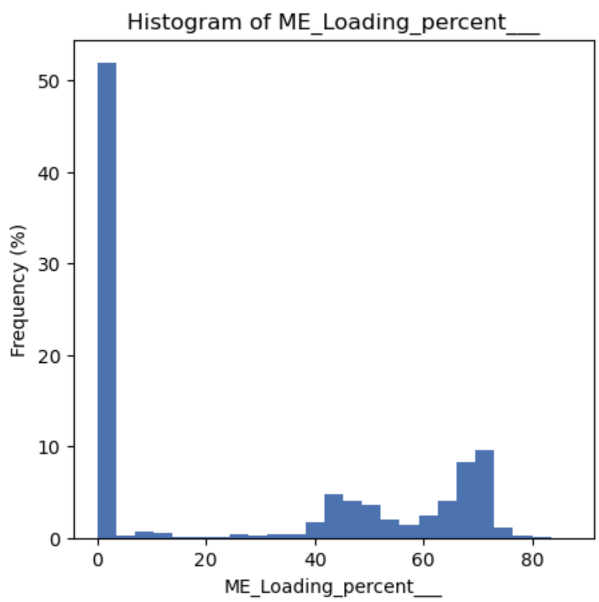
Figure 8: Time series plots representing the merged dataset's basic parameters.


Histogram of Speed_Over_Ground_knots_
Frequency (%)
0
10
20
30
40
50
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
Speed_Over_Ground_knots_


Histogram of Speed_Through_Water_knots_
Frequency (%)
0
5
10
15
20
25
30
−5
0
5
10
15
Speed_Through_Water_knots_


Histogram of MERPM_AMS_rpm_
Frequency (%)
0
10
20
30
40
50
−80
−60
−40
−20
0
20
40
60
80
MERPM_AMS_rpm_


Histogram of Propeller_Shaft_Power_kW_
Frequency (%)
0
10
20
30
40
50
0
2000
4000
6000
8000
Propeller_Shaft_Power_kW_


Histogram of ME_Loading_percent___
Frequency (%)
0
10
20
30
40
50
0
20
40
60
80
ME_Loading_percent___

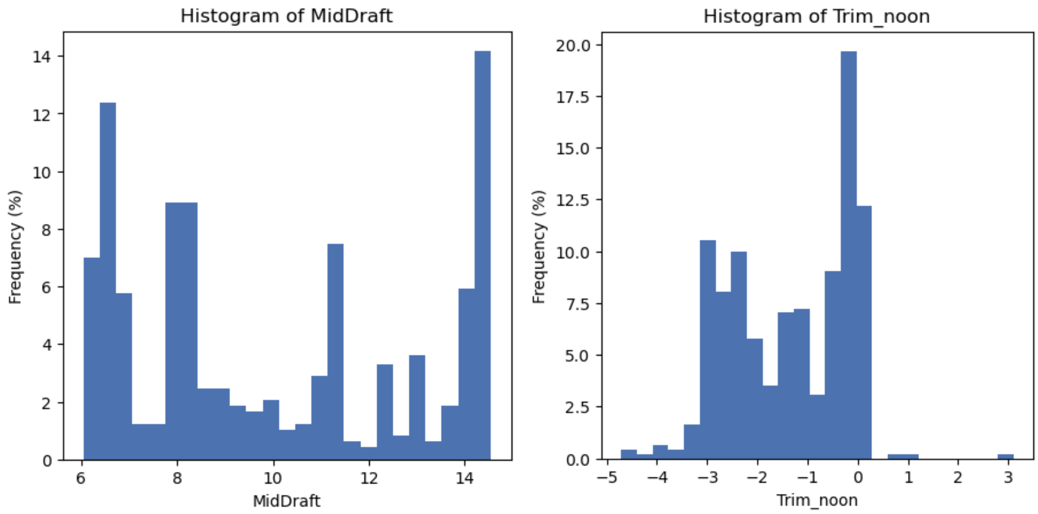


Figure 9: Histograms representing the basic parameters of the merged Dataset.

To complete the profiling of the final merged data set, it is crucial to present the parameters that were created due to deleting several parameters as well as the merge of the two data sets. Hence, the following table was created:

Parameters
Speed Over Ground (SOG) [knots]
Speed Through Water (STW) [knots]
Propeller Shaft Power (PSP) [kW]
ME Revolutions per Minute [rpm]
ME Loading percent [%]
Fuel Index Position [n/a]
Vessel Heading [deg]
Relative Wind Direction [deg]
Relative Wind Speed [m/sec]
Fuel Oil Temperature (ME return) [^oC]
Fuel Oil Temperature (ME supply) [^oC]
Shaft Torque [kNm]
Shaft Thrust [kN]
M/E Shaft Revolutions [rpm]
Ballast Condition
Cargo Carried [tn]
Fuel-LCV [kJ/kg]
Fore Draft [m]
Mid Draft [m]
Aft Draft [m]
Heading
Air Temperature
Sea Temperature
Sea Height
Swell
Swell Height
Wind Force
Trim

Table 3: Parameters of the merged data set.

3.2 Data Filtering

3.2.1 Threshold Values

3.2.1.1 Speed

One of the most important operating parameters, if not the most important, is the ship's speed which is described by the Speed through water (STW) and Speed over ground (SOG) parameters. Its monitoring is essential to the shipowner as it is highly correlated with the ship's fuel consumption. Additionally, speed is tracked through sensors that compute the absolute values of the aforementioned parameters, hence both parameters should have only positive values. Lastly, low water speed readings are also related to a ship's approach to, operation within, or departure from a port. Some situations are not covered by this study since there is either no fuel oil usage or very little. Therefore, the following threshold values were applied to exclude points associated with sensor failure as well as port operation:

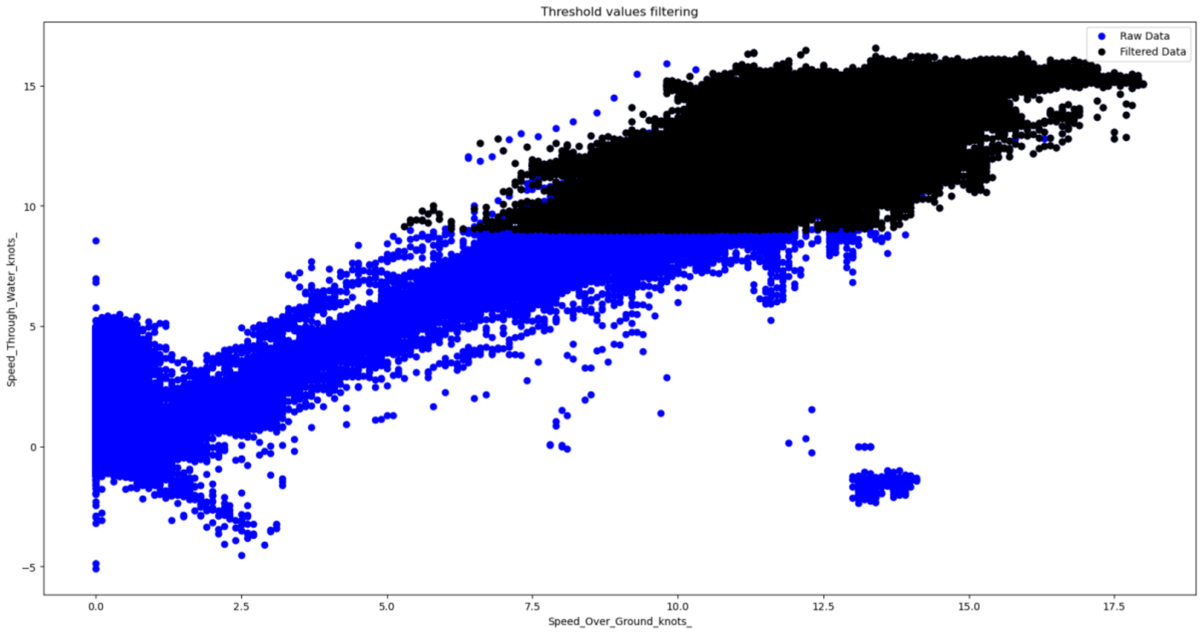
- $\boldsymbol{STW > 9\,(kn)}$

3.2.1.2 ME Revolutions per Minute (ME RPM)

The ME revolutions are highly correlated to the Propeller Shaft Power parameter since both of them are connected to the engine. Similarly, to the Vessel's Speed Through Water parameter, low ME revolutions as well as Shaft Power values, reflect the vessel's operation inside port terminals. Hence, in order to exclude points associated with the port operation the following threshold values were applied:

- $\boldsymbol{ME\ RPM > 50\,(rpm)}$
- $\boldsymbol{PSP > 3000\,(kW)}$

The following plots (Figure 10) reflect the application of the abovementioned filters (Speed, ME RPM & PSP):



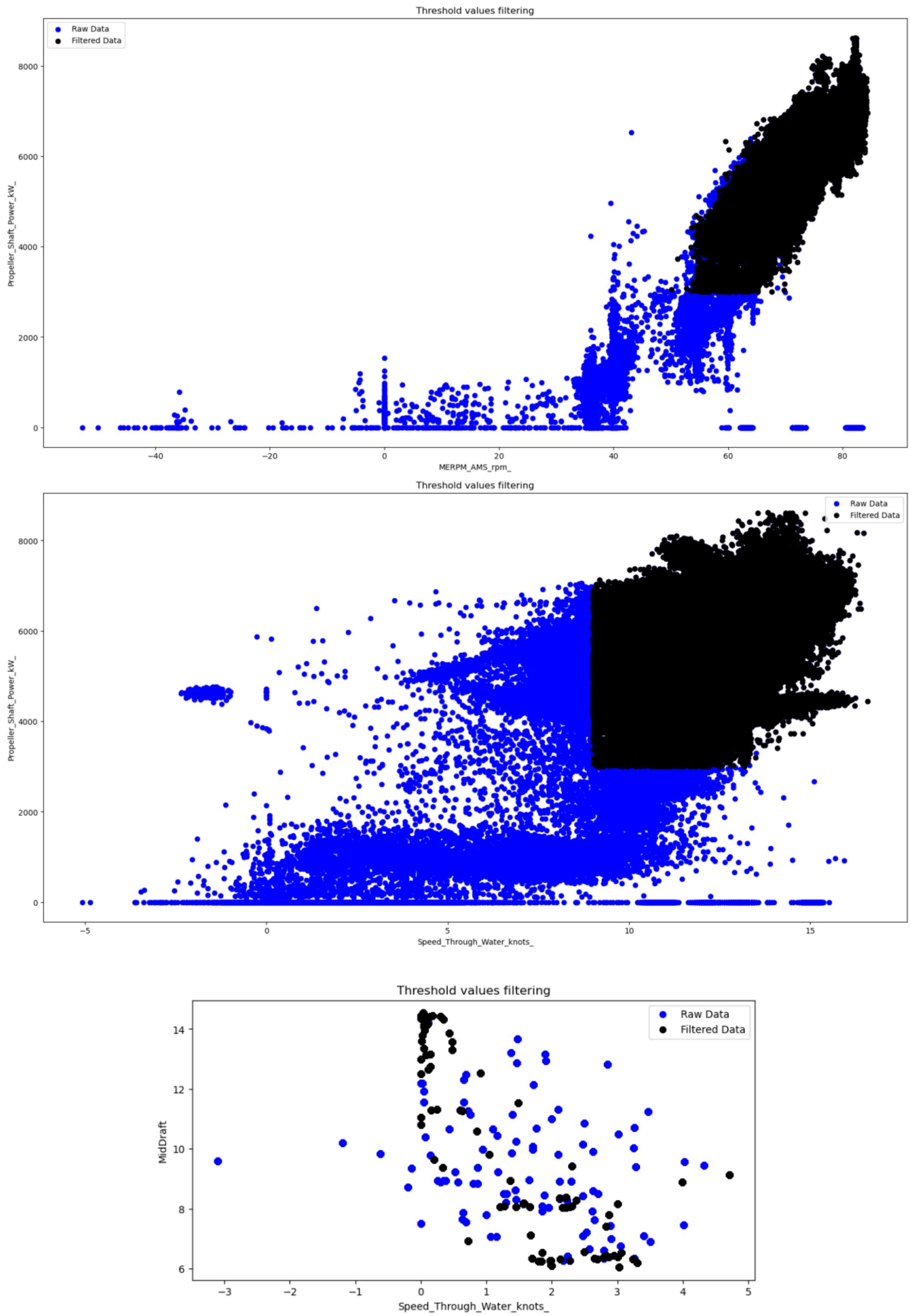


Figure 10: Plots reflecting the application of the threshold values.

3.3 Data Cleaning

3.3.1 Null Values

As mentioned before the given data set has many values that need to be excluded from the model development. At this phase, any discrepancies, mistakes, or missing numbers are cleaned up and processed out of the data. This calls for addressing outliers, coping with missing data, and getting rid of duplicates. From Table 2 it was observed that some sensors were not working as they should, therefore it was concluded that values coming from these sensors should be eliminated. Hence, *DTN AIR TEMPERATURE 10M ACTUAL (ºC), DTN AIR PRESSURE MEAN SEA LEVEL ACTUAL (mbar), DTN SEA TEMPERATURE 0M ACTUAL (ºC), DTN SIGNIFICANT WAVE HEIGHT (m), DTN MEAN WAVE DIRECTION (deg) and ME Fuel Oil Consumption (mt/day), Water Depth Relative to the Transducer_BRG_ECHO (m)* parameters were excluded from the data set. Lastly, after the filtering of the values the remaining null values were the following:

Parameter	**Missing Values**
TIME	0
Speed Over Ground (knots)	22
Speed Through Water (knots)	0
Propeller Shaft Power (kW)	0
ME Revolutions per Minute (rpm)	0
ME Loading (%)	1492
Fuel Index Position	54
Vessel Heading (deg)	9
Relative Wind Direction (deg)	12
Relative Wind Speed (m/sec)	16
FO Temperature at ME Return (Cº)	4
FO Temperature at ME Supply (Cº)	4
Shaft Torque (kNm)	4
Shaft Thrust (kN)	1
Ballast Condition	337
Cargo carried (tons)	19
Fuel LCV (kJ/kg)	1649
Fwd Draft (m)	0
Mid Draft (m)	0
Aft Draft (m)	0
Heading	0
Air Temperature	0
Sea Temperature	0
Sea Height	0
Swell	0
Swell Height	0
Wind Force	0
Trim	0

Table 4: Null values of filtered data.

At this phase, it is needed to drop the zero values from our data set since they are likely to cause problems in the development of the predictive model. After this step we are left with the following data set:

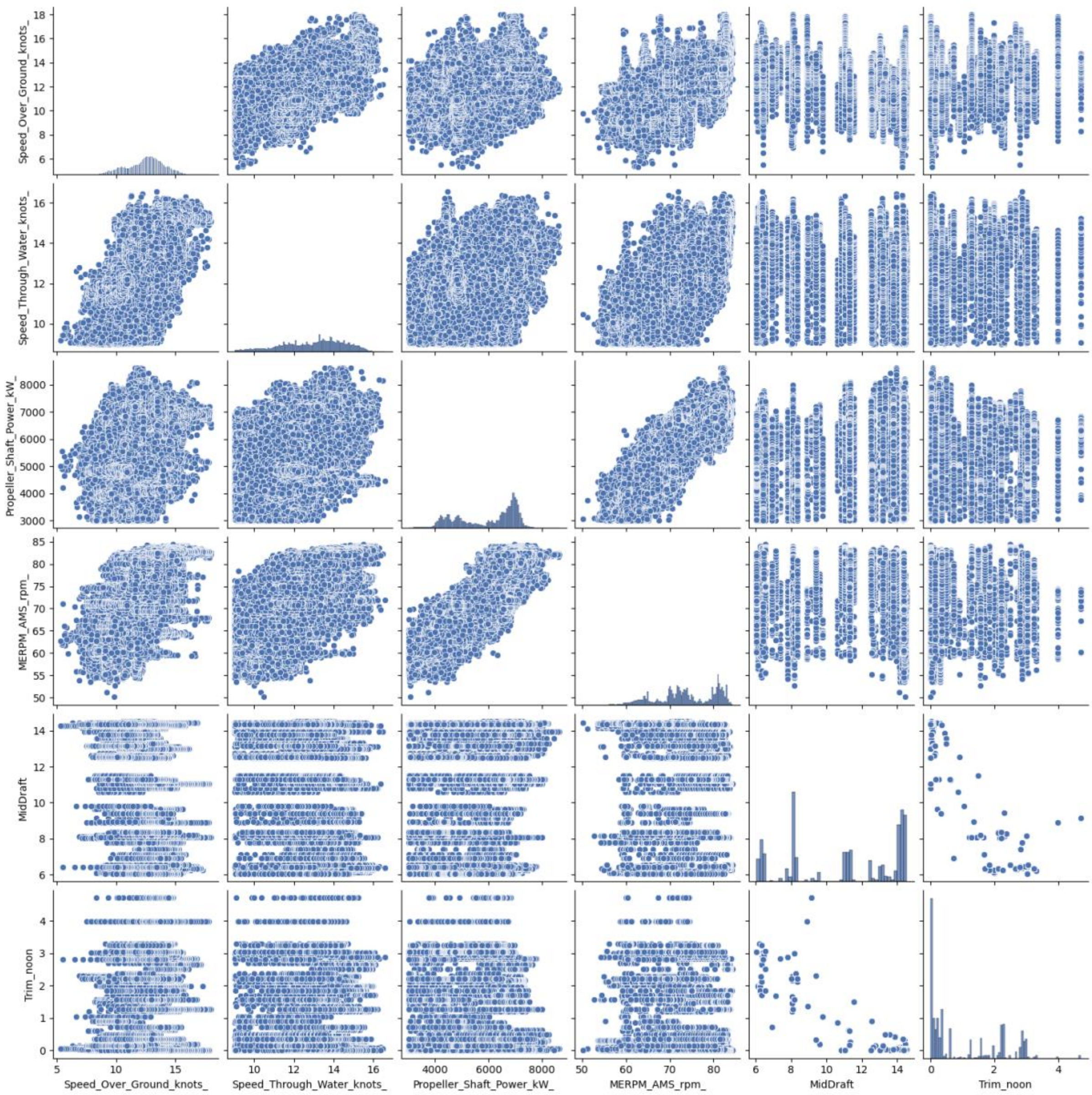


Figure 11: Plots of filtered data without null values.

3.4 Parameters correlation

3.4.1 Pearson Correlation Coefficient

Before detecting outliers, through a statistical outlier detection method, it is necessary to investigate the correlation between the dataset's parameters. There are several ways to do that, although in data analysis the most popular is the Pearson Correlation Coefficient method.

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 * \sum(y_i - \bar{y})^2}} \tag{3-1}$$

Where,

$$r = correlation\ coefficient$$
$$x_i = values\ of\ the\ x - variable\ in\ a\ sample$$
$$\bar{x} = mean\ of\ the\ values\ of\ the\ x - variable$$
$$y_i = values\ of\ the\ y - variable\ in\ a\ sample$$
$$\bar{y} = mean\ of\ the\ values\ of\ the\ y - variable$$

The status, direction, and strength of the association between the two variables are all determined using correlation analysis, a statistical technique that aids in explaining the relationship between variables. The determination of the Pearson Correlation Coefficient (r) is crucial for the analysis because the Pearson method is employed in correlation analysis. The range of the r coefficient is from -1 to +1. Positive correlation and negative correlation are denoted by plus and minus signs in front of the coefficient, respectively. One of the critical factors to consider is the measure of the connection between Propeller Shaft Power (PSP) and other inputs. This is because the correlation coefficient can help determine the relative significance or weight of various factors that affect the predicted PSP. The **Error! Reference source not found.** reveals that the vessel speed and shaft parameter indicators are strongly correlated with the propeller shaft power indicator. These inputs can be prioritized based on their correlation values.

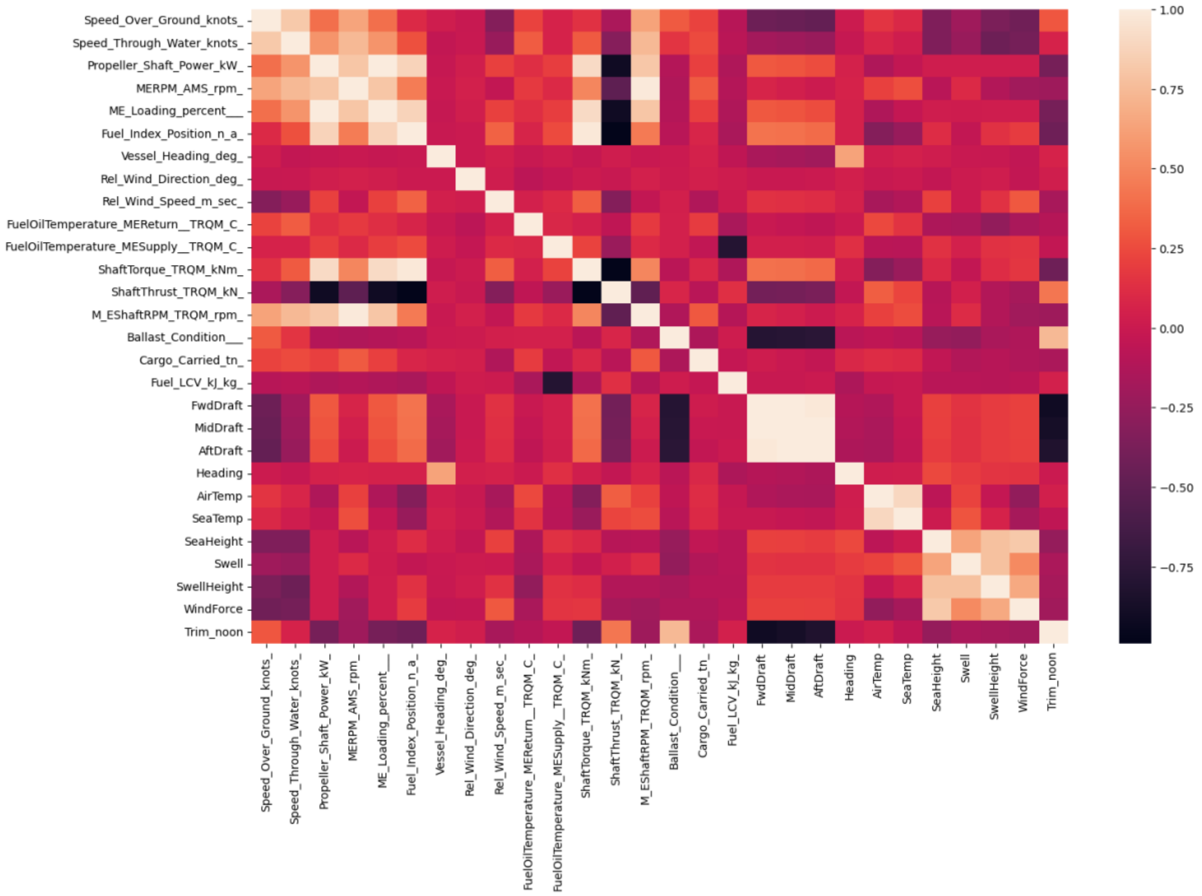


Figure 12: Pearson Correlation heatmap.

3.4.2 Parametric Plots

Despite the parameter correlation calculated by the Pearson Correlation Coefficient method, having an initial estimation of the anticipated correlation between these parameters is crucial, drawing from the governing principles of the phenomena in which they are implicated. These laws of physics are described through the following mathematical equations:

- Regarding the engine's operation:

$$P_{en} = Q \cdot 2\pi \cdot n_{en} \qquad [3\text{-}2]$$

Where, P_{en} is the engine's power outcome (Break horsepower / BHP), Q is the crankshaft's torque and n_{en} are the revolutions per second of the engine.

- The empirical Propeller Law:

$$P_{prop} = c \cdot V^3 \qquad [3\text{-}3]$$

The empirical propeller law is a relationship that describes the power consumed by a ship's propeller. It is an empirical equation derived from observations and experimental data. The propeller law states that the power consumed by the propeller, denoted as P_{prop}, is proportional to the cube of the ship's speed, V.

Alternatively, the propeller law can also be written in terms of the propeller revolutions, denoted as n, which are often proportional to the engine's revolutions. In this form, the equation becomes:

$$P_{prop} = c \cdot n^3 \quad [3\text{-}4]$$

The constant c in both equations represents the efficiency and characteristics of the propeller system under specific operating conditions.

The empirical propeller law provides valuable insights into the power requirements and performance characteristics of ship propellers. By understanding this relationship, ship designers and operators can make informed decisions regarding propeller selection, optimization, and overall vessel performance.

- The Calm Water Resistance Coefficient:

$$C_T = \frac{R}{\frac{1}{2}\rho S V^2} \quad [3\text{-}5]$$

Here, R represents the measured resistance force, ρ represents the density of the fluid, and S represents the wetted surface area. The water resistance refers to the force that must be overcome by the propeller's thrust in calm sea conditions to achieve the desired speed V. This is achieved when the propeller's effective power (Peff) is equal to V·R.

To compare the laws of physics with the data captured, the following plots were created. In general, it is observed that they comply with the basic parametric plots provided by the aforementioned mathematical equations.

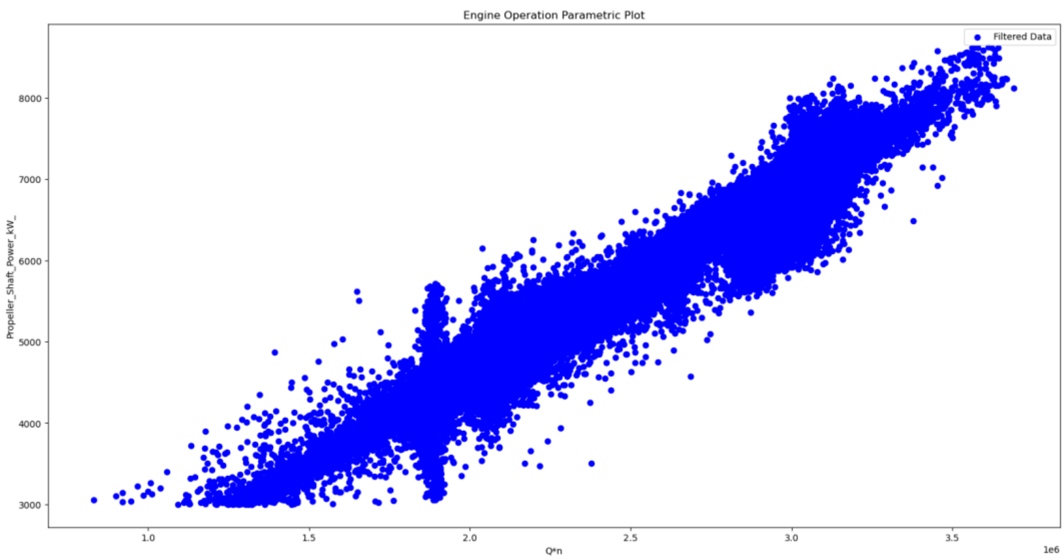


Figure 13: Engine's Operation Equation visualization.

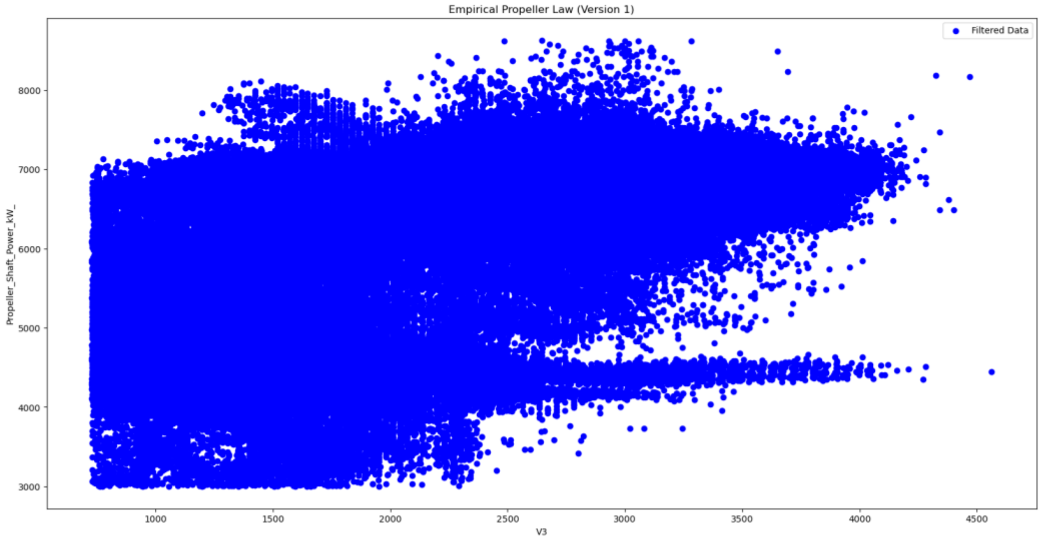


Figure 14: Empirical Propeller Law (STW Version)

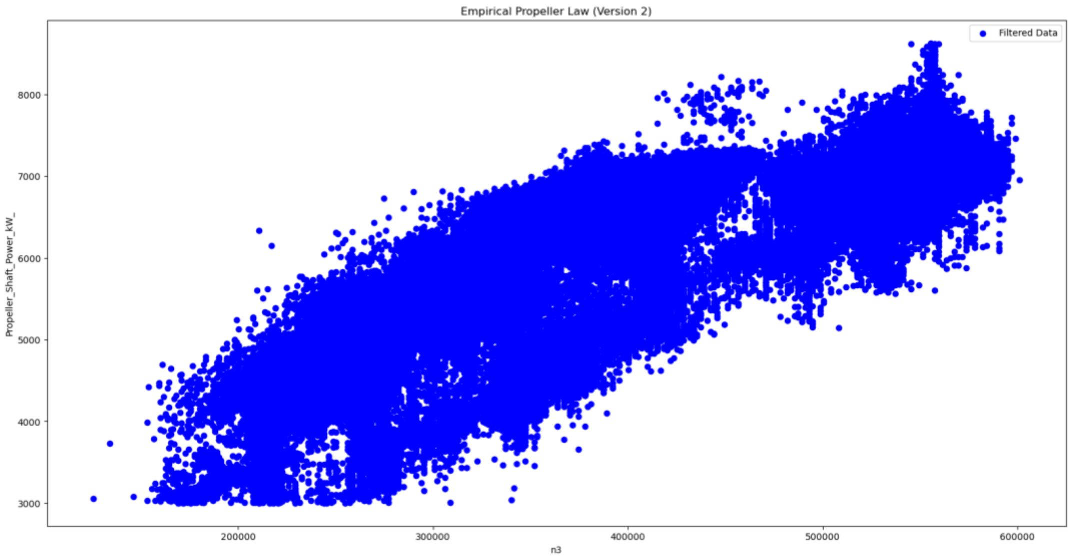


Figure 15: Empirical Propeller Law (RPM Version)

After observing the empirical propeller law parametric plot (Figure 15: Empirical Propeller Law (RPM Version)), it became apparent that the merged dataset consisted of two distinct clusters. This distinction was expected as the vessel underwent maintenance during the recorded period. In particular, on 19/07/2021, a propeller repair was conducted, which significantly impacted all fundamental propulsion parameters. In order to visually analyse the variation in vessel performance efficiency, it became essential to partition the dataset into two portions: one representing data before the maintenance and the other after.

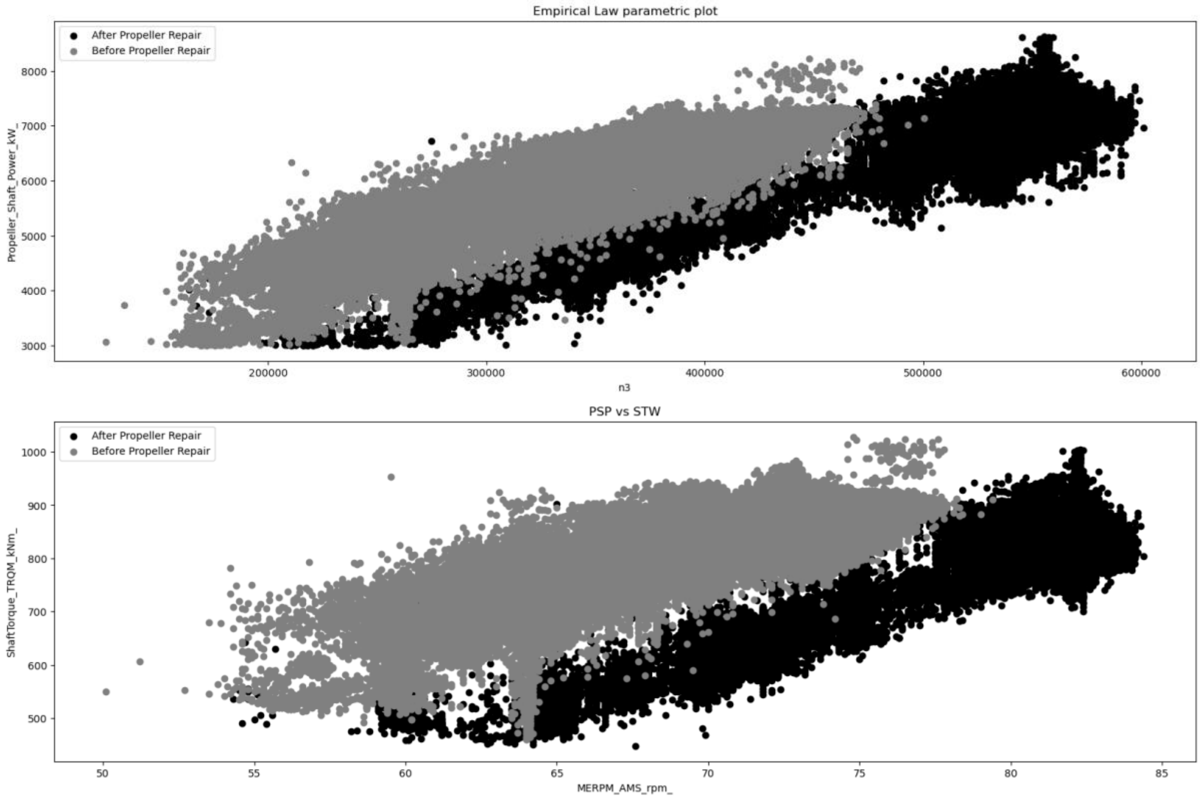


Figure 16: Visual observation of propeller repair effect on propulsion efficiency.

Ultimately, Figure 16: Visual observation of propeller repair effect on propulsion efficiency. confirms our hypothesis of the cluster existence and show the effect of the propeller repair on the efficiency of the vessel's performance. It is clear that after the repair the performance efficiency level has risen since both $\frac{PSP}{n^3}$ and $\frac{TRQM}{n}$ ratios have decreased.

3.5 Outlier Detection

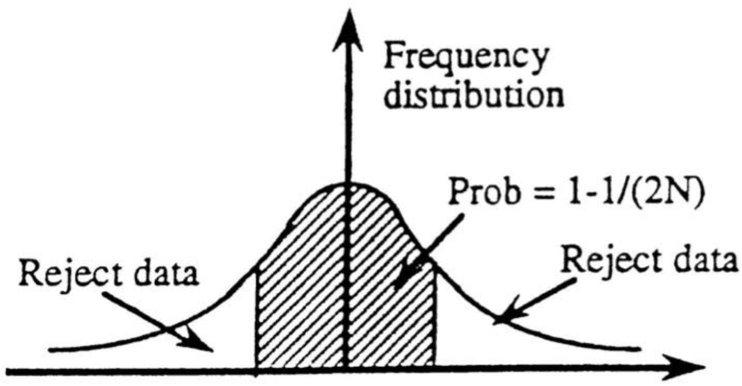
Outlier detection is a critical step in data analysis as it serves multiple purposes. By identifying and handling outliers, we ensure the quality and integrity of the data, preventing errors that can arise from data collection or recording. Accurate analysis is achieved by mitigating the influence of outliers on statistical calculations and models, ensuring reliable and meaningful results. Outlier detection also enhances the robustness of algorithms and models by mitigating the disproportionate impact of extreme values. Additionally, outliers can offer valuable insights and patterns, contributing to a deeper understanding of the data and potential discoveries. Ultimately, outlier detection facilitates informed decision-making processes, enabling actions based on reliable and accurate information.

Additionally, Outlier detection is connected to, yet separate from, noise removal and noise accommodation. While all three concepts address undesired noise in data, they have distinct objectives. Noise refers to irrelevant elements in the data that hinder analysis. Noise removal focuses on eliminating these unwanted components prior to analysis. On the other hand, noise accommodation aims to protect statistical model estimation from anomalous observations, effectively shielding the model from their influence. (Singh & Upadhyaya, 2012)

To detect the outlying and noisy data, an alternative method was used which is mostly based on the Chauvenet's criterion, a statistical method that identifies outliers through the following steps:

1. Calculation of the mean value of the samples: $\mu = \frac{1}{N}\sum_i^N d_i$ [3-6]
2. Calculate the standardized deviation: $delta_i = |(d_i - \mu)|$ [3-7]
3. Calculation of the standard deviation of the samples: $\sigma = \sqrt{\frac{1}{N}\sum_i^N delta_i}$ [3-8]
4. Calculation of the probability for the occurrence of any value d_i: $P(d_i) = erfc(\frac{delta_i}{\sigma\sqrt{2}})$ [3-9]
5. A sample is considered an outlier if the following inequality is fulfilled: $P(d_i) \cdot N < 0.5$ [3-10]

(Rochim, 2016)



(*Statistical Rejection of "Bad" Data-Chauvenet's Criterion*, n.d.)
Figure 17: Chauvenet's criterion

Combining Chauvenet's criterion along with the study of (P. Karagiannidis, 2019), the filtering procedure developed for the detection of outliers is described in the following steps:

1. Select a primary parameter X whose values are to be filtered.
2. Split the primary parameter X in groups of values with range v.
3. Select a secondary parameter Y which is highly correlated with the primary parameter X.
4. For each group Gi of X, normalize Y with z-score normalization.
5. Select an outlier threshold k, where k ∈ [2, 3.5].
6. For every respective value of Y in the Gi group, Yij , test if the following inequality is fulfilled:

$$|Yij| \leq k \quad [3\text{-}11]$$

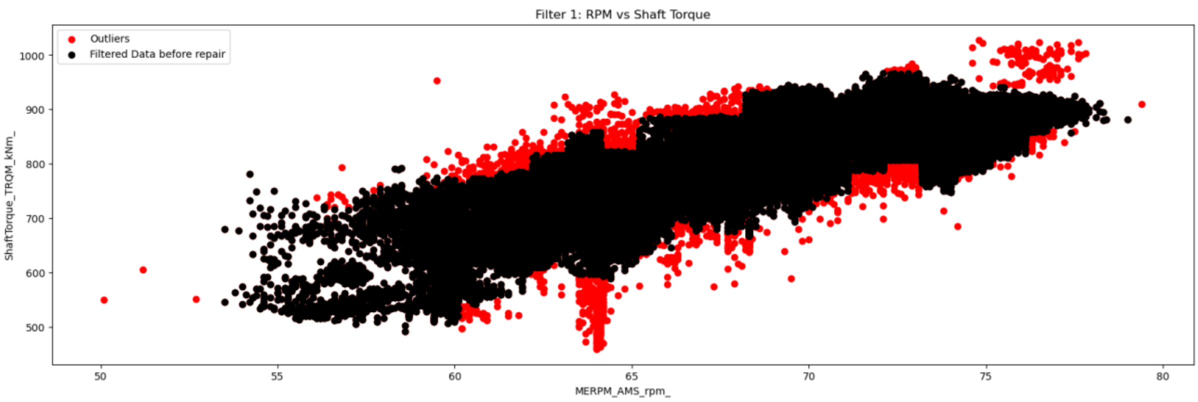
7. If the inequality is not fulfilled, reject the data point.

For the purpose of this study, the primary parameters X were selected to be the ME revolutions (ME RPM) which was split into groups of values with a range of 1 rpm, and SOG which was split into groups of values with a range of 0.5 knots, while the secondary parameters were chosen, according to Figure 12: Pearson Correlation heatmap., to be the Propeller Shaft Torque, the Propeller Shaft Power (PSP), and the Speed Through Water (STW). Finally, the threshold for all three parameters was set to k=2.5. To observe the outcome of the outlier filtering process, the following plots were created:

3.5.1 Filter 1: Shaft Torque – RPM

Primary Parameter	RPM
Secondary Parameter	Shaft Torque
k	2.5
v (rpm)	1
data points removed (%)	1.63 %

Table 5: Filter 1 details



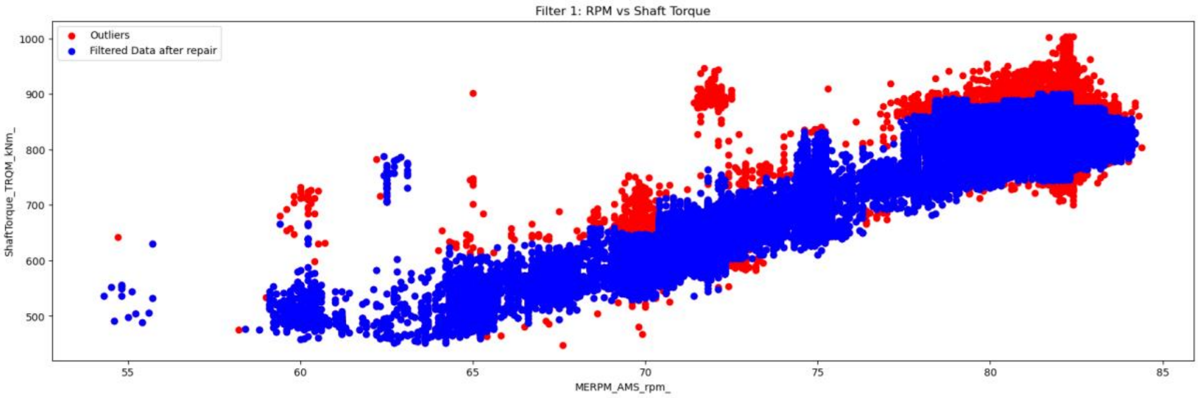


Figure 18: Plot visualization of statistical outliers of the propeller's shaft torque & ME revolutions.

3.5.2 Filter 2: PSP – RPM

Primary Parameter	RPM
Secondary Parameter	PSP
k	2.5
v (rpm)	1
data points removed (%)	0.65 %

Table 6: Filter 2 details.

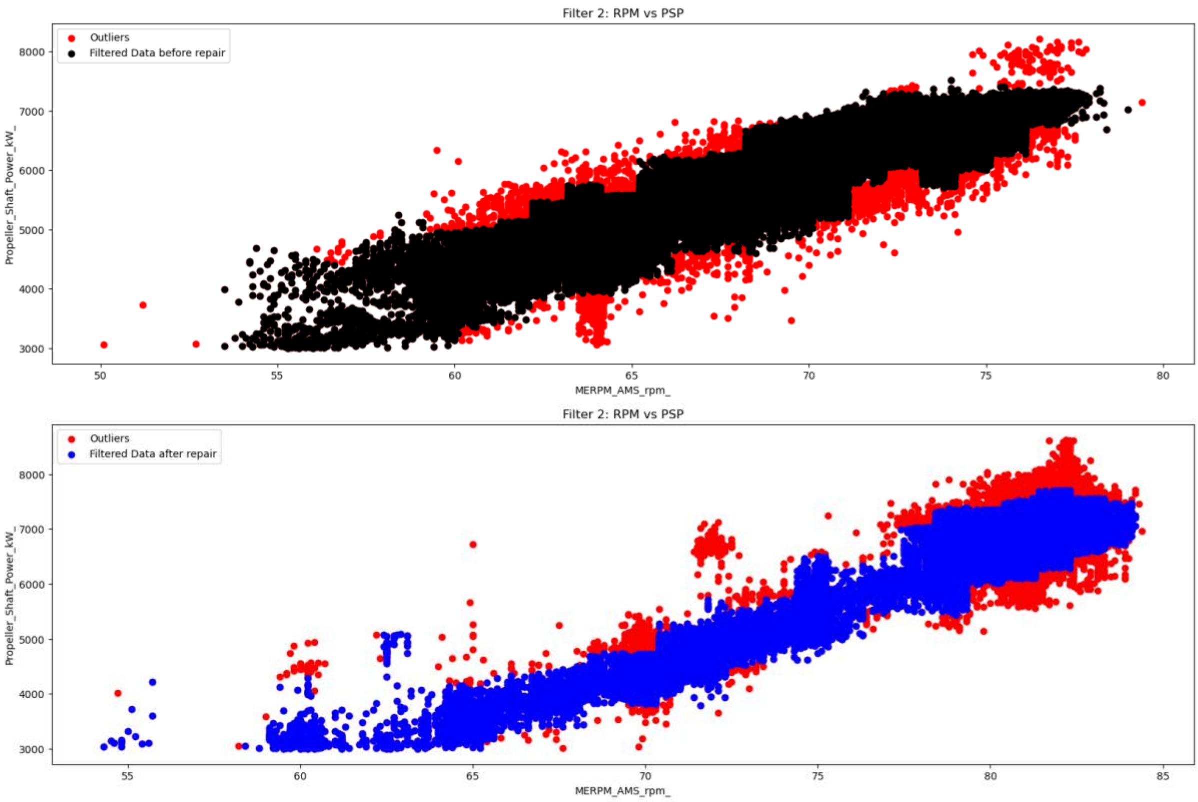


Figure 19: Plot visualization of statistical outliers of the ME revolutions & PSP.

3.5.3 Filter 3: STW – RPM

Primary Parameter	RPM
Secondary Parameter	STW
k	2.5
v (rpm)	1
data points removed (%)	2.04 %

Table 7: Filter 3 details.

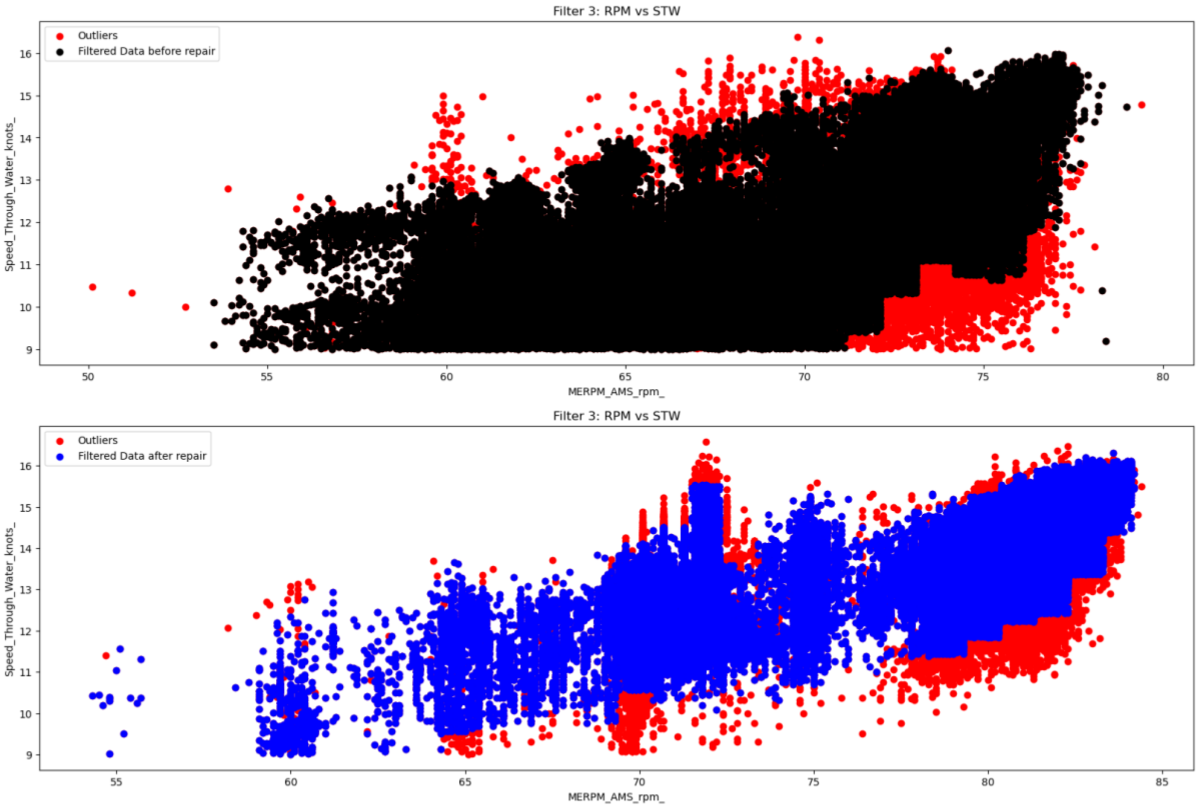


Figure 20: Plot visualization of statistical outliers of the ME revolutions & Speed Through Water.

3.5.4 Filter 4: STW – SOG

Primary Parameter	SOG
Secondary Parameter	STW
k	2.5
v (kn)	0.5
data points removed (%)	2.16 %

Table 8: Filter 4 details.

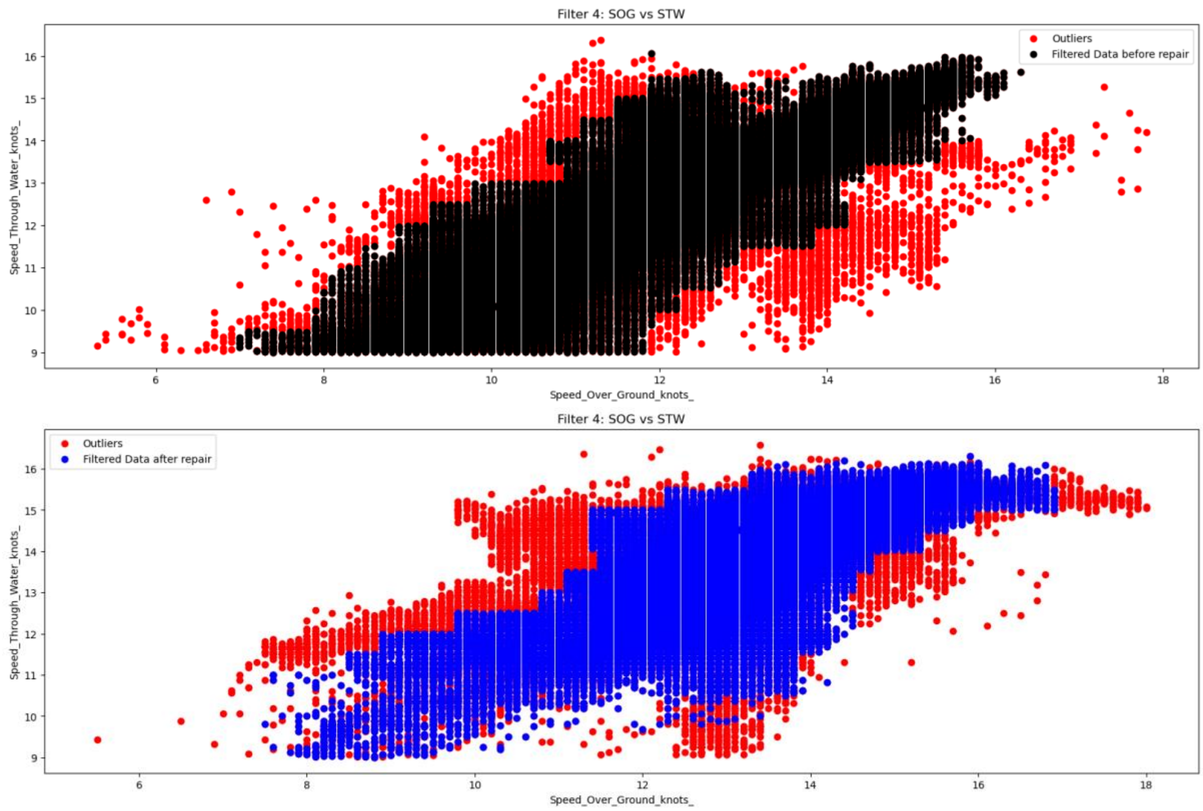


Figure 21: Plot visualization of statistical outliers of the Speed Over Ground & Speed Through Water.

3.6 Data Smoothing

Data smoothing is a methodology used in data analysis to reduce noise and detect underlying trends or patterns in data. The purpose of this technique is to create a more refined version of the original data set, by applying a mathematical algorithm or function. Some of the techniques employed in data smoothing include moving averages, exponential smoothing, and kernel smoothing. Moving averages utilize the computation of the average of neighboring data points, whereas exponential smoothing gives more weight to recent data points. In kernel smoothing, a specified kernel function is used to fit a curve to the data. Data smoothing can be useful in revealing trends and patterns that may not be immediately noticeable from the raw data. Furthermore, it can aid in eliminating noise or anomalies from a data set, thus simplifying analysis and interpretation. Nonetheless, applying data smoothing techniques should be made with extreme caution since they can sometimes obscure crucial data details or introduce artifacts or biases.

In this study, a simple moving average smoothing method was used. A simple moving average (SMA) is an arithmetic moving average calculated by adding all the previous *n* values and then dividing that figure by the number of time periods in the calculation average. Since the data were recorded with a frequency of 1 minute, the period number of the values will define the time window that will be used. For the purpose of this study a 15-minute window was selected.

$$SMA = \frac{A_1 + A_2 + \cdots + A_n}{n} \quad [3\text{-}12]$$

$$A_n: The\ arithmetic\ value\ of\ the\ variable\ at\ the\ period\ n$$

$$n: The\ number\ of\ total\ periods$$

To evaluate the effect of smoothing, the Speed-Through-Water parameter was plotted over a 2000-sample period.

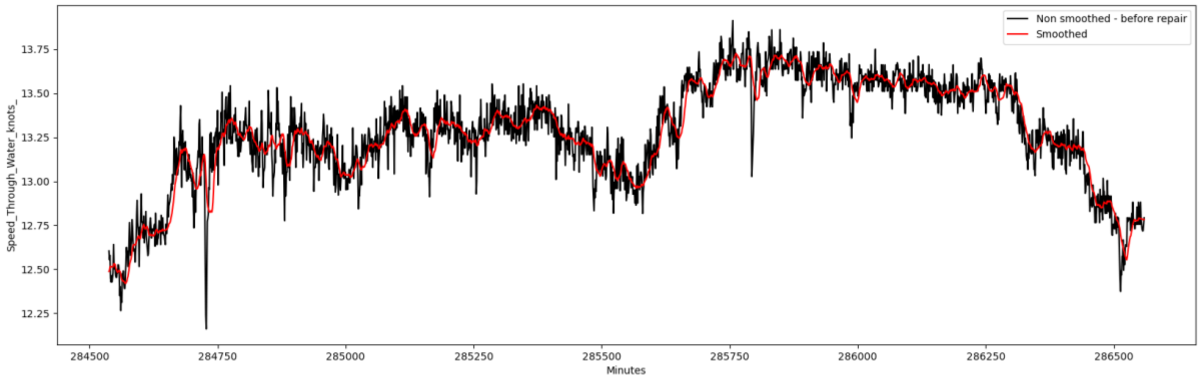


Figure 22: Smoothed – Non-Smoothed data curve comparison (before repair Dataset)

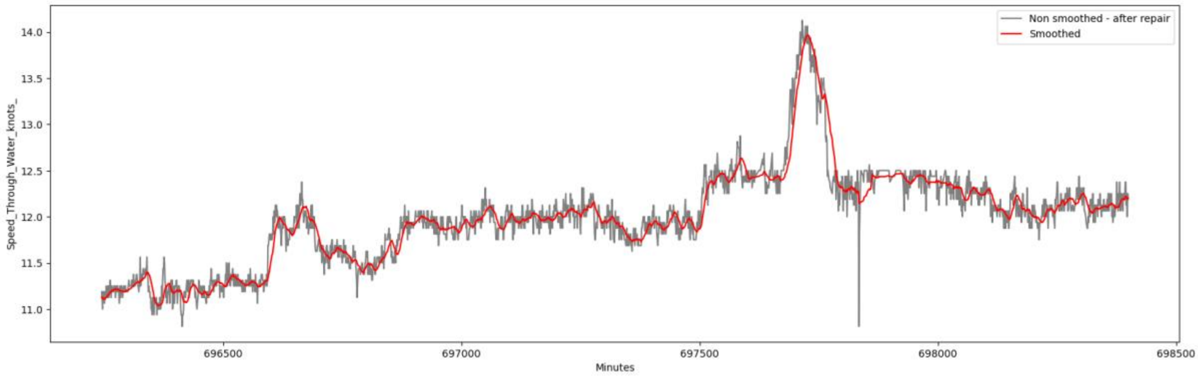


Figure 23:Smoothed – Non-Smoothed data curve comparison (after repair Dataset)

From the above graphs, it is recognizable that the smoothed data curve, in both datasets, is far more stable.

3.7 Data Quality Check

In data science, data quality checks are of utmost importance for reliable analysis, valid conclusions, and optimal model performance. The type of insights derived from predictive models along with their accuracy, heavily rely on high-quality data. By identifying and addressing issues such as missing values, outliers, inconsistencies, and inaccuracies through data quality checks, they ensure that the data used for analysis is reliable and trustworthy. Hence, it is essential to dedicate a chapter to the evaluation of the pre-processing procedure of the data.

$$Dataset\ 1: Raw\ data\ containing\ both\ noon\ and\ telemetry\ data.$$

$$Dataset\ 2: Dataset\ 1\ without\ the\ threshold\ values.$$

$$Dataset\ 3: Dataset\ 2\ without\ null\ values.$$

$$Dataset\ 4: Dataset\ 3\ without\ statistical\ outliers.$$

$$Dataset\ 5: Final\ Dataset\ |\ Smoothed\ Dataset\ 4.$$

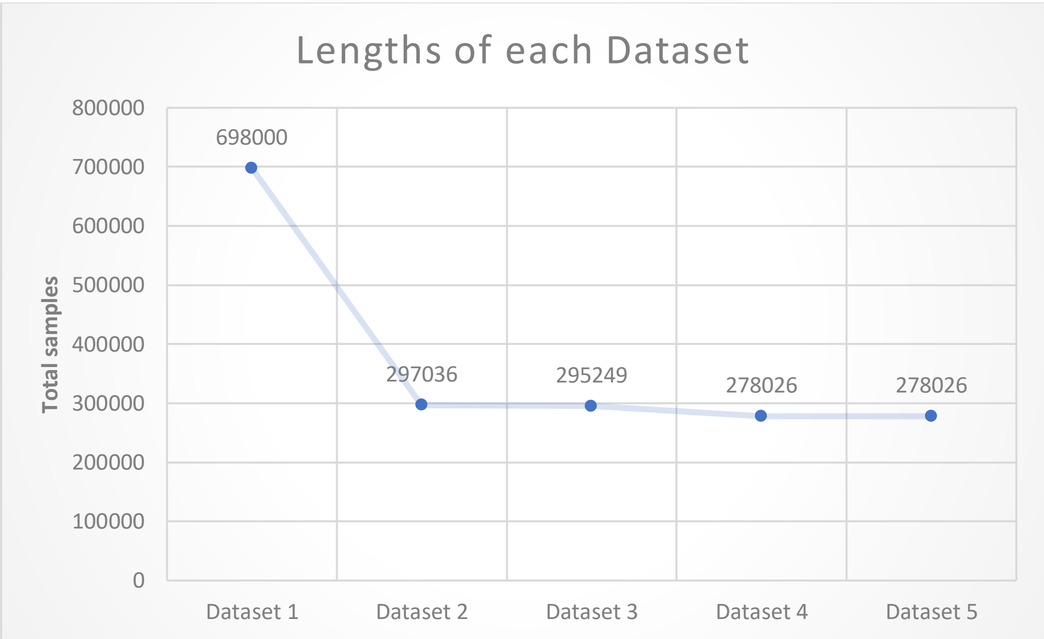


Figure 24: Total points of each dataset.

The quality of data is of major importance to data scientists since it impacts both the validity and the performance of machine learning predictive models and data science projects. To ensure that our data is of high-quality and reflect the physics and mathematical laws of ship propulsion, data quality checks were implemented by measuring statistical parameters and comparing the following plots.

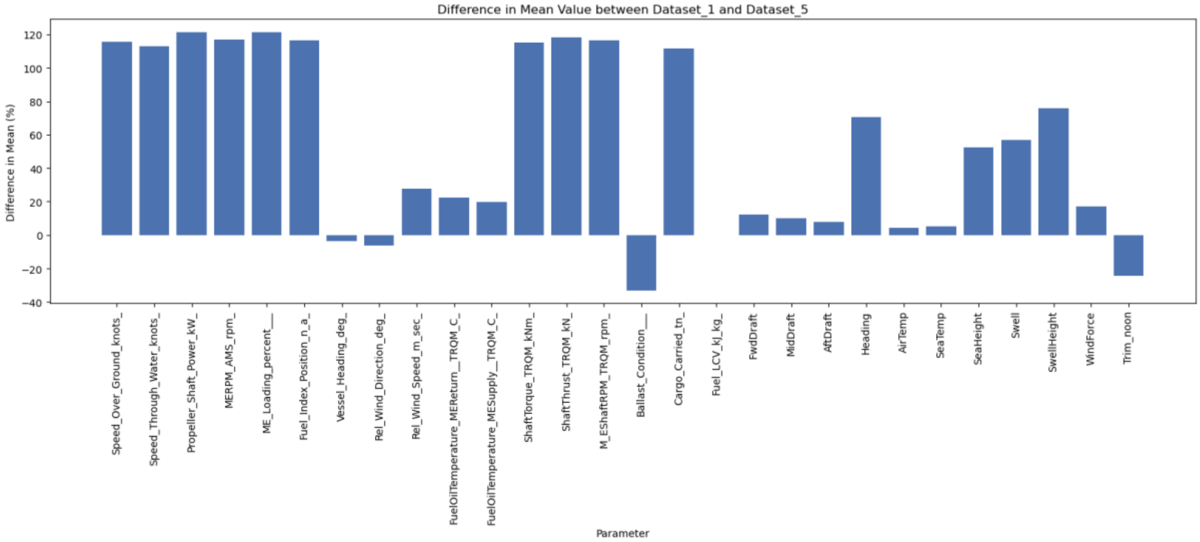


Figure 25: Difference in Mean value between Dataset1 and Dataset5

From Figure 25 it is easy to notice that the mean values of the raw dataset have changed significantly due to the data filtering. Although, it would be wiser to evaluate this change for every dataset and compare the influence of each pre-processing method.

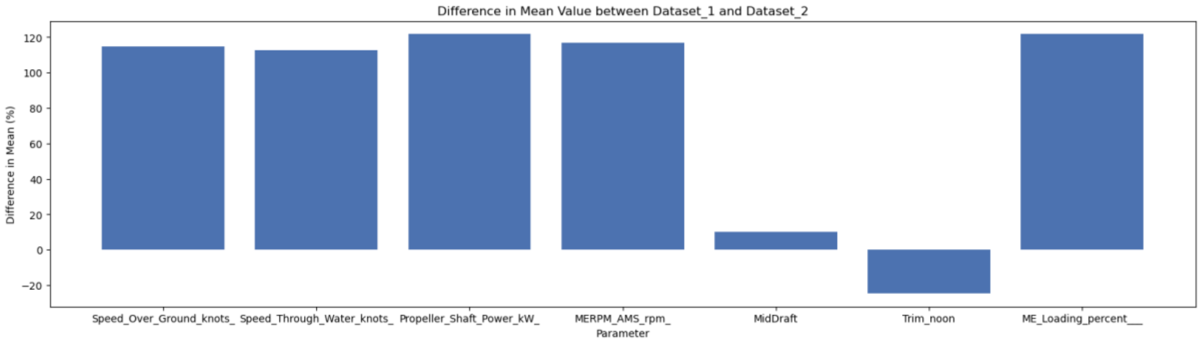


Figure 26: Effect of threshold filtering in the parameter's mean values.

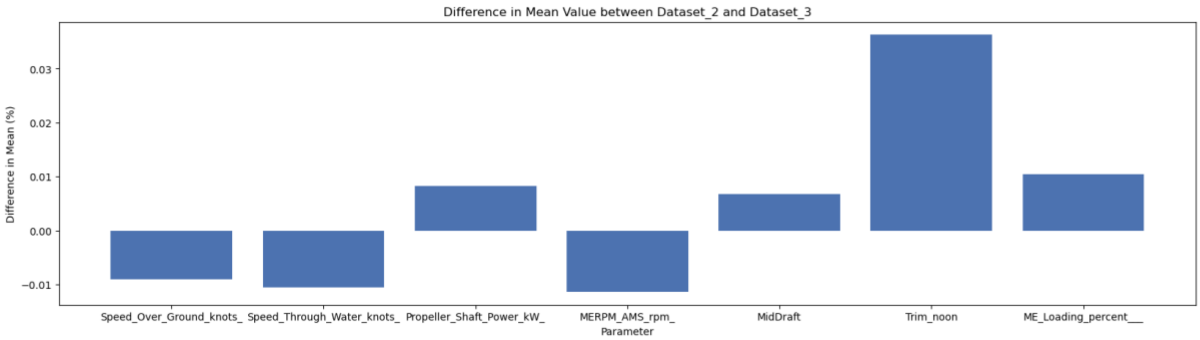


Figure 27: Effect of null values dropping in the parameter's mean values.

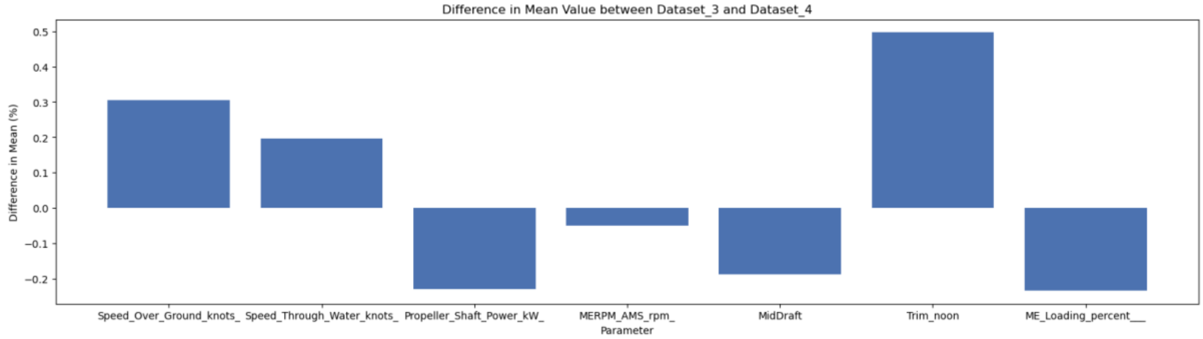


Figure 28: Effect of outlier's detection on the parameter's mean values.

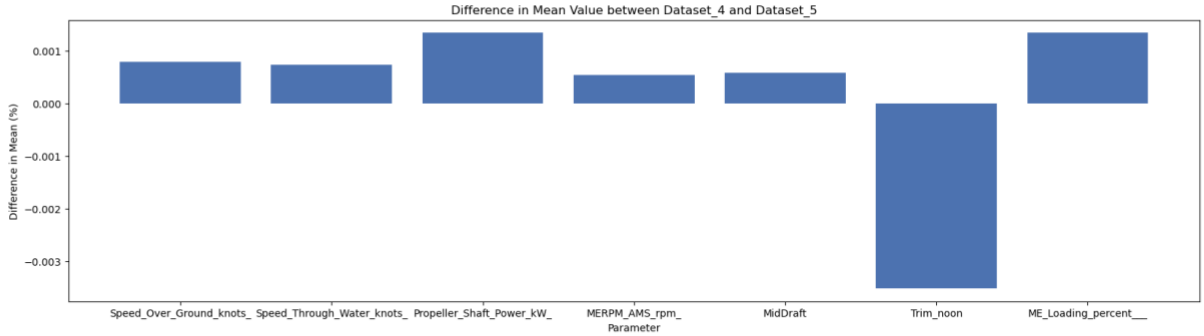
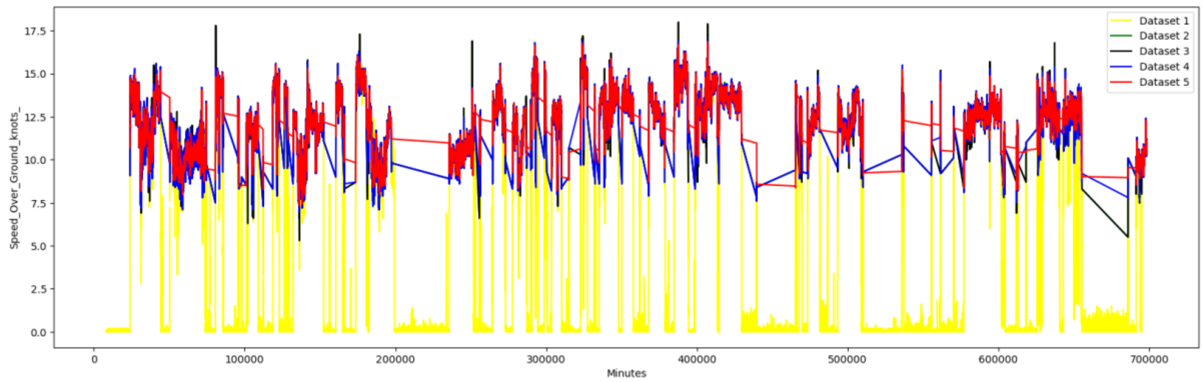


Figure 29: Effect of data smoothing on the parameter's mean values.

The difference between the raw and final dataset's mean values arises mainly from the threshold values filtering procedure, while the next steps do not affect the dataset's values as much, but they succeed in maintaining a stable Dataset.

Furthermore, poor-quality data can introduce biases, noise, or skewed representations, resulting in suboptimal model performance and inaccurate predictions. Data quality checks help identify and address these issues, improving the accuracy and robustness of the models. By ensuring that the data used to train and test the models is of high quality, data scientists can enhance the performance and reliability of their machine learning models. The following plots examine the differences between each dataset in time giving us the ability to evaluate the Final Dataset.



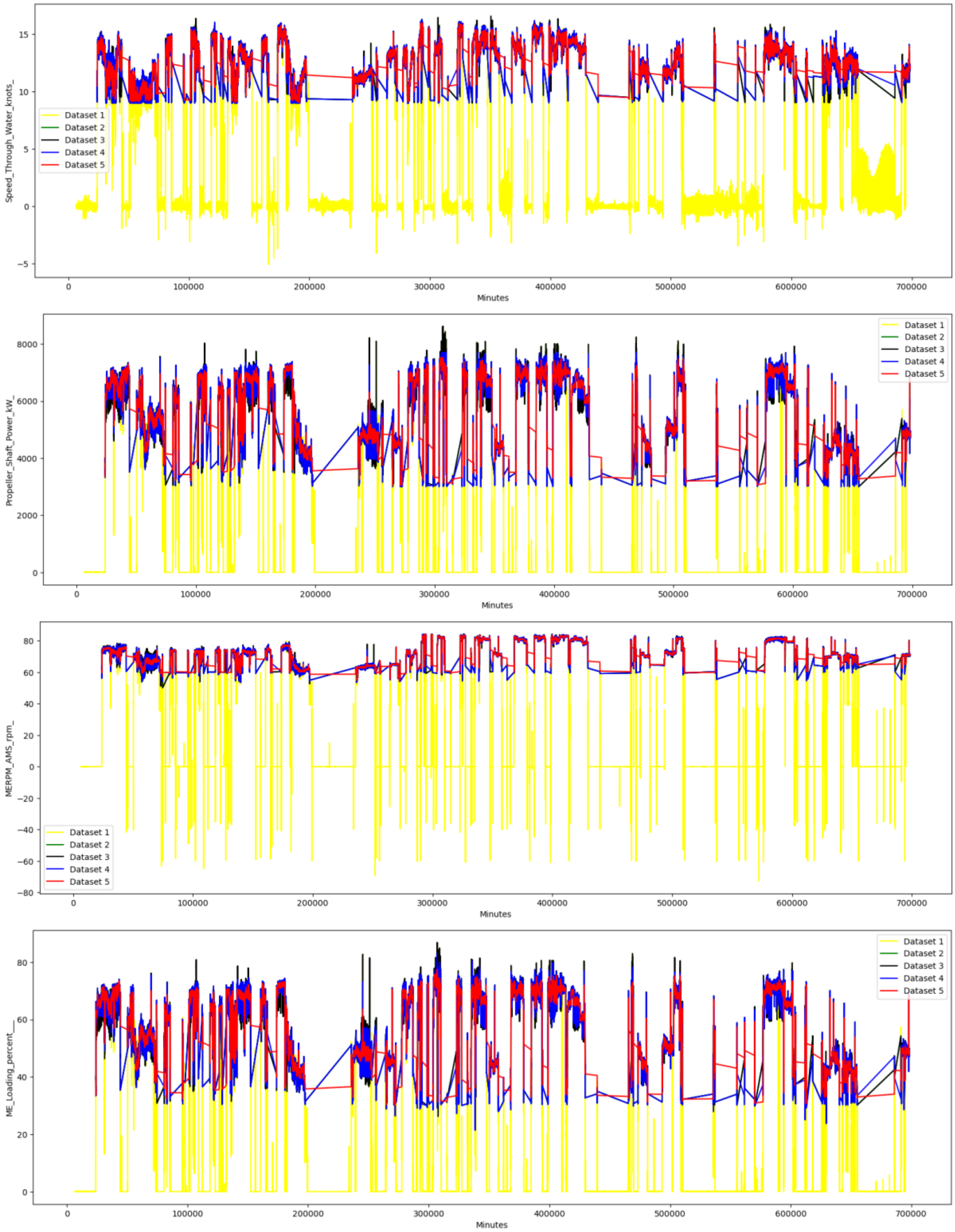


Figure 30: Plots in time visualizing the difference between the values of each dataset.

In summary, data quality checks are vital for data scientists as they enable reliable analysis, ensure valid conclusions, and enhance model performance. By addressing issues such as missing values, outliers, inconsistencies, and inaccuracies, data scientists can work with reliable data, have confidence in their findings, and build accurate and robust models. Ultimately, data quality checks contribute to the overall success and effectiveness of data science projects.

3.8 Final Dataset

After evaluating the quality of the data we've managed to create, it is essential to present the final Dataset which will lead us to accurate predictions and valid accusations. Hence, this chapter is to appraise the final dataset and observe its statistical characteristics.

	Speed_Over_Ground_knots_	Speed_Through_Water_knots_	Propeller_Shaft_Power_kW_	MERPM_AMS_rpm_	MidDraft	Trim_noon	ME_Loading_percent___
count	122364.000000	122364.000000	122364.000000	122364.000000	122364.000000	122364.000000	122364.000000
mean	11.839911	12.368038	5813.282278	69.227105	11.328924	-1.052820	58.540438
std	1.665272	1.654435	1029.369704	5.085015	3.274407	1.165401	10.369137
min	7.320000	9.057820	3046.186667	54.856667	6.100000	-4.710000	30.109427
25%	10.480000	11.080493	4829.173333	64.346667	8.070000	-2.220000	48.630002
50%	11.833333	12.389157	6059.946667	70.140000	14.080000	-0.300000	61.030933
75%	13.066667	13.739565	6786.986667	73.713333	14.300000	-0.060000	68.348300
max	16.026667	15.798253	7302.400000	77.160000	14.540000	0.000000	73.538780

Table 9: Descriptive statistics for Final Dataset – before repair.

	Speed_Over_Ground_knots_	Speed_Through_Water_knots_	Propeller_Shaft_Power_kW_	MERPM_AMS_rpm_	MidDraft	Trim_noon	ME_Loading_percent___
count	155634.000000	155634.000000	155634.000000	155634.000000	155634.000000	155634.000000	155634.000000
mean	13.015700	13.504561	6005.658851	77.426471	10.634575	-0.981172	60.476002
std	1.185177	1.180046	1163.019327	5.222399	2.843533	1.096862	11.712186
min	8.213333	9.171560	3047.573333	59.420000	6.050000	-3.240000	30.629340
25%	12.393333	12.632938	4826.880000	71.860000	8.080000	-1.840000	48.609073
50%	13.026667	13.608333	6582.613333	80.120000	11.260000	-0.350000	66.242880
75%	13.700000	14.401166	6981.546667	81.480000	13.000000	-0.060000	70.303340
max	16.846667	16.004120	7530.240000	84.100000	14.430000	0.000000	77.487480

Table 10: Descriptive statistics for Final Dataset – after repair.

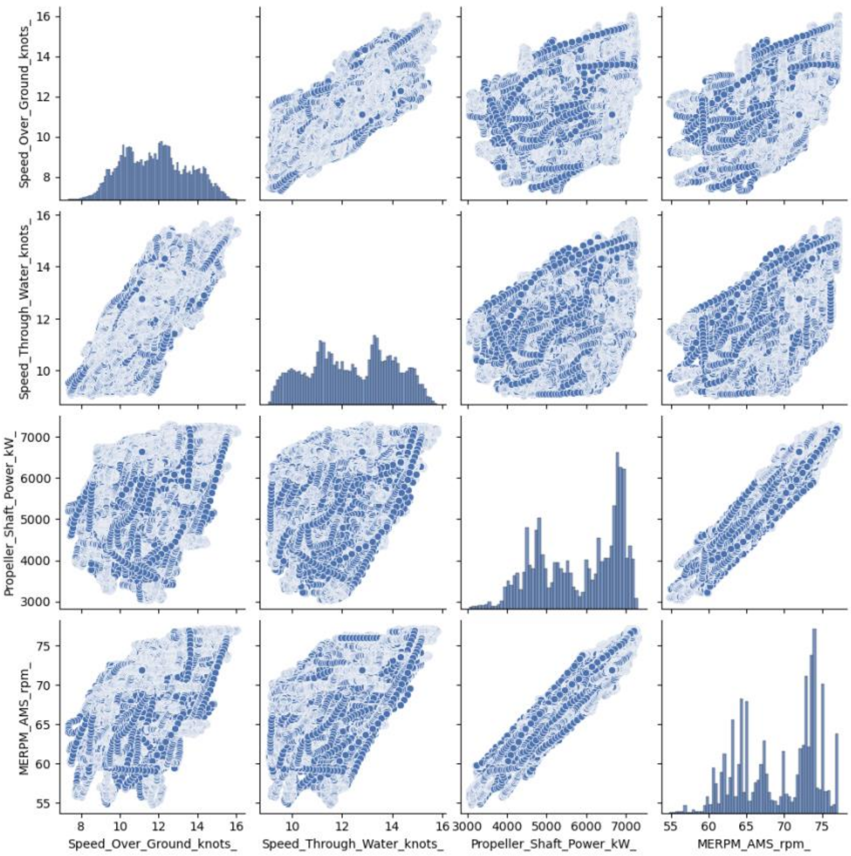


Figure 31: Main features of the final data set – before repair.

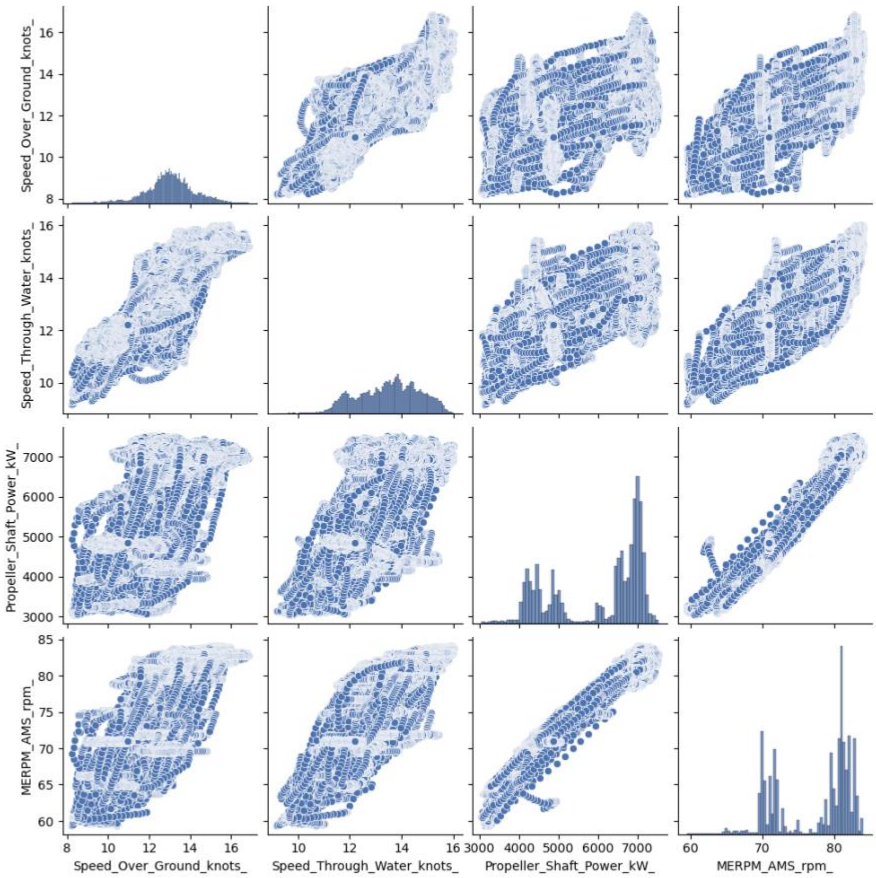


Figure 32: Main features of the final data set – after repair.

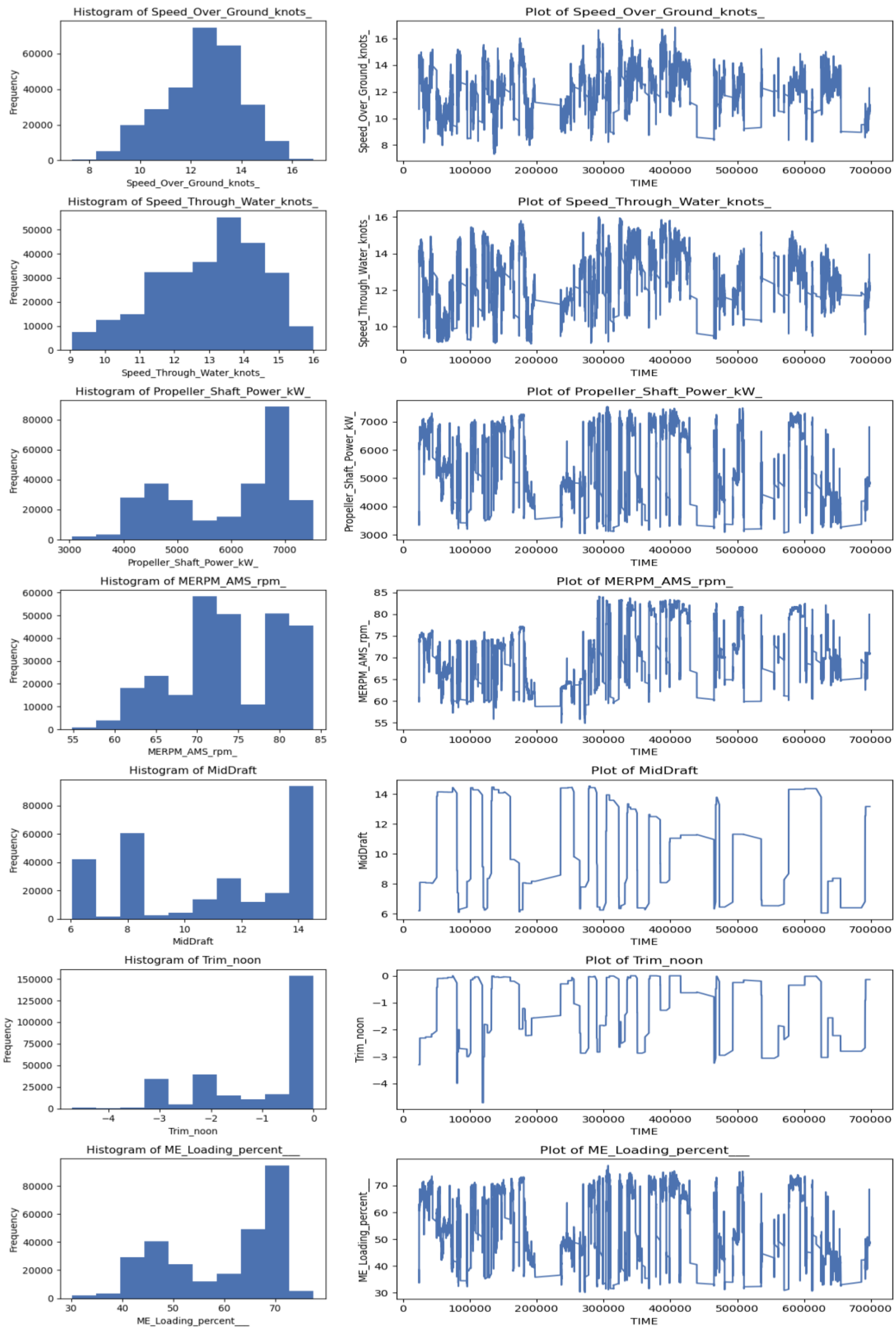


Figure 33:Histograms & Time series plots for Final Dataset's important parameters.

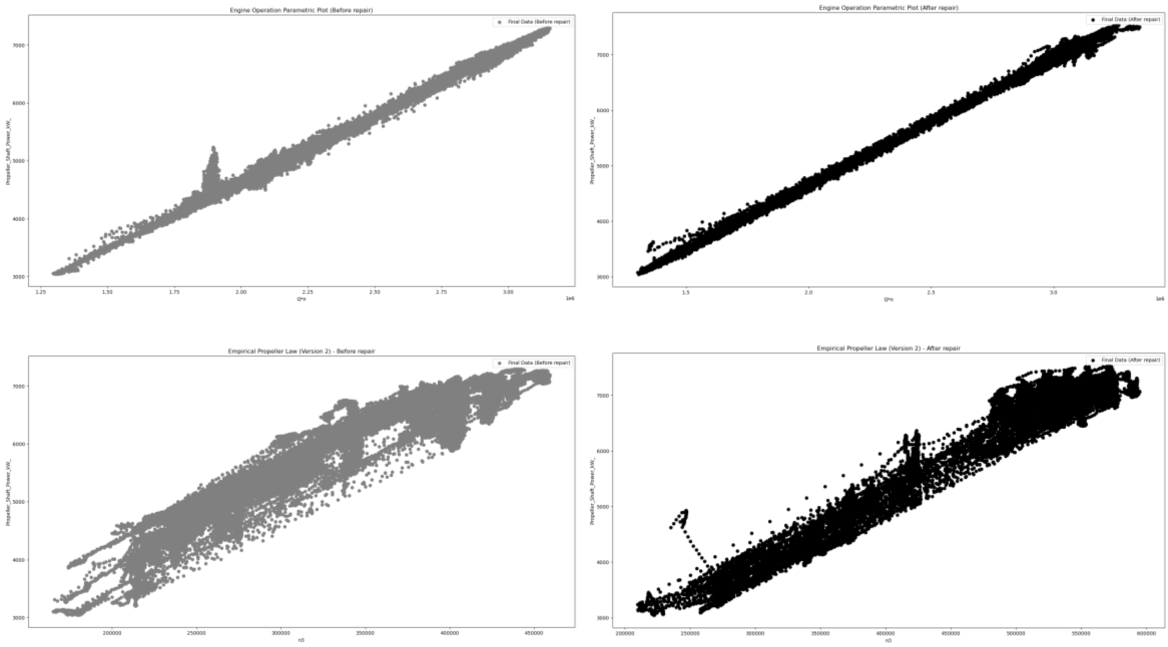


Figure 34:Parametric plots of the Final Dataset (Before & After repair).

Finally, based on the visualization of the final datasets plot and the statistical characteristics, it is observed that the propeller repair which happened during the captured period, had a significant influence on the vessel's operation since after comparing Table 9 and Table 10 it was noticeable that the vessel achieved higher values of Speed through water while delivering approximately the same amount of shaft power.

4 Feature Engineering

The action of choosing, modifying, and converting unprocessed data into features that can be applied in supervised learning is known as feature engineering. Feature engineering through the years has evolved to be a fundamental part of machine learning science, since it showed its ability to create and train better features to improve machine learning efficiency and expand into a wider variety of assignments.

The accuracy and generalization performance of a machine learning model can be significantly influenced by the quality and relevance of the features used for training. Feature engineering is a technique utilized in machine learning that uses data to generate new variables that are not originally included in the training set. This approach can create fresh features for both supervised and unsupervised learning, aiming to revolutionize and accelerate data processing while improving the model's accuracy.

The process of feature engineering includes creating, transforming, extracting, and selecting variables, also known as features, aiming to build an accurate machine-learning algorithm. The following are the main processes involved in feature engineering:

Feature Creation: The procedure of discovering the variables that are most important for the predictive model. It's a subjective procedure that depends on the developer's creativity and expertise. To create new, more powerful derived features, existing features are joined using addition, subtraction, multiplication, and division.

Transformations: To enhance the performance of a model, transformations require changing the predictor variables. Ensuring the model can handle a range of data, scaling variables to the same range to make the model easier to understand, increasing accuracy, and preventing computational errors by making sure all features are within the model's acceptable range are all examples of transformations made to improve the model's performance.

Feature Extraction: Extraction of features from raw data entails automatically creating new variables. Automatically condensing the volume of data into a more manageable collection for modeling is the purpose of this stage. Techniques for feature extraction include principal component analysis, edge detection algorithms, text analytics, and cluster analysis.

Feature Selection: Feature selection algorithms examine, rank, and assess different features to determine which ones are necessary for the model and should be prioritized, which ones are redundant and should be removed, and which ones are unimportant and should be eliminated.

4.1 Feature Creation

From the data set provided to perform this study, the values measured through various sensors, mentioned above, are used as the first group of features. Although, through the data preprocessing these parameters were evaluated and some of them were eliminated. Therefore, the group of features created by the final data set consists of the following parameters:

Parameters
Time
Speed Over Ground (knots)
Speed Through Water (knots)
Propeller Shaft Power (kW)
ME RPM (rpm)
ME Loading percentage (%)
Fuel Index Position
Vessel Heading (deg)
Relative Wind Direction (deg)
Relative Wind Speed (m/sec)
Fuel Oil Temperature ME Return TRQM
Fuel Oil Temperature ME Supply TRQM
Shaft Torque TRQM (kNm)
Shaft Thrust (kN)
ME Shaft RPM TRQM (rpm)
Ballast Condition
Cargo Carried (tons)
Fuel LCV (kJ/kg)
Fore Draft (m)
Mid Draft (m)
Aft Draft (m)
Heading
Air Temperature
Sea Temperature
Sea Height
Swell
Swell Height
Wind Force
Trim

Table 11: Available parameters of final dataset after preprocessing.

4.2 Feature Extraction

4.2.1 Wind Effect

To be able to extract information about the wind effect on the vessel, it is required to transform the relative wind direction units from degrees to radians. After transforming the units of wind direction, the next step is to extract a wind effect feature through the Relative Wind Direction (rad) and Relative Wind Speed (m/sec) parameters. Hence, the following steps were followed:

1. Apply the cosine function to the Relative Wind Direction (rad) feature, in order to normalize the data and constrain it from -1 to 1.
2. Multiply the cosine of the Relative Wind Direction feature with the Relative Wind Speed (m/sec) to get the new feature that resembles the effect of wind forces on the vessel. ("Wind Effect").

$$Wind\ Effect = \cos(Rel.Wind\ Dir.) * [Rel.Wind\ Speed] \qquad [4\text{-}1]$$

4.2.2 Currents

Water currents refer to the movement of water in a particular direction. They can have a significant impact on the vessel's performance since they can influence its speed and fuel consumption. Hence, it was deemed important to create a feature that describes the current speed of water. To do that, the Speed Through Water was subtracted from the Speed Over Ground.

$$Currents\ (m/sec) = SOG - STW \qquad [4\text{-}2]$$

4.2.3 Power Output

To generate the parametric plots (Figure 34) the following parameters were generated which resemble the power output of the vessel's engine.

- $Q * n = \text{Shaft Thrust (kN)} * \text{ME RPM (rpm)} * 60$ [4-3]
- $V^3 = (Speed\ Through\ Water\ (m/sec))^3$ [4-4]
- $n^3 = (ME\ RPM\ (rpm))^3$ [4-5]

4.2.4 Days Since Repair (DSR)

This feature was created as an indication of the propeller maintenance importance as well as an indicator of fouling. From the dataset before the repair, we can observe how as the days gone by the performance of the vessel decreases, while from the dataset after the repair it is possible to observe the effect of the maintenance. Hence the following feature was created:

$$DSR = \text{Date} - \text{Repair Date} \quad [4\text{-}6]$$

Where Repair Date was set to be equal to 19/07/2021.

4.2.5 Mean Draft

The mean draft feature is created to replace the Fore and Aft Drafts and represent them as a unique value.

$$Mean\ Draft = \frac{Aft\ Draft + Fore\ Draft}{2} \quad [4\text{-}7]$$

4.3 Feature Selection

The performance of the prediction model depends on the selection of the features. Therefore, a broad search domain for selecting distinct features is used in this study to ensure that there is no overfitting problem while enhancing prediction performance. While the official group of parameters consists of 32 variables, the algorithms investigated in this study are easy to overfit due to their dimensionality, therefore it is necessary to remove intercorrelated and unnecessary variables which would not have any contribution to the improvement of the model's performance. In order to determine which features are to be selected, Correlation analysis (Pearson coefficient) and Random Forest Regression model were used.

4.3.1 Correlation Analysis

To determine the correlation between the input variables and the output (PSP) a Pearson correlation matrix heatmap was created.

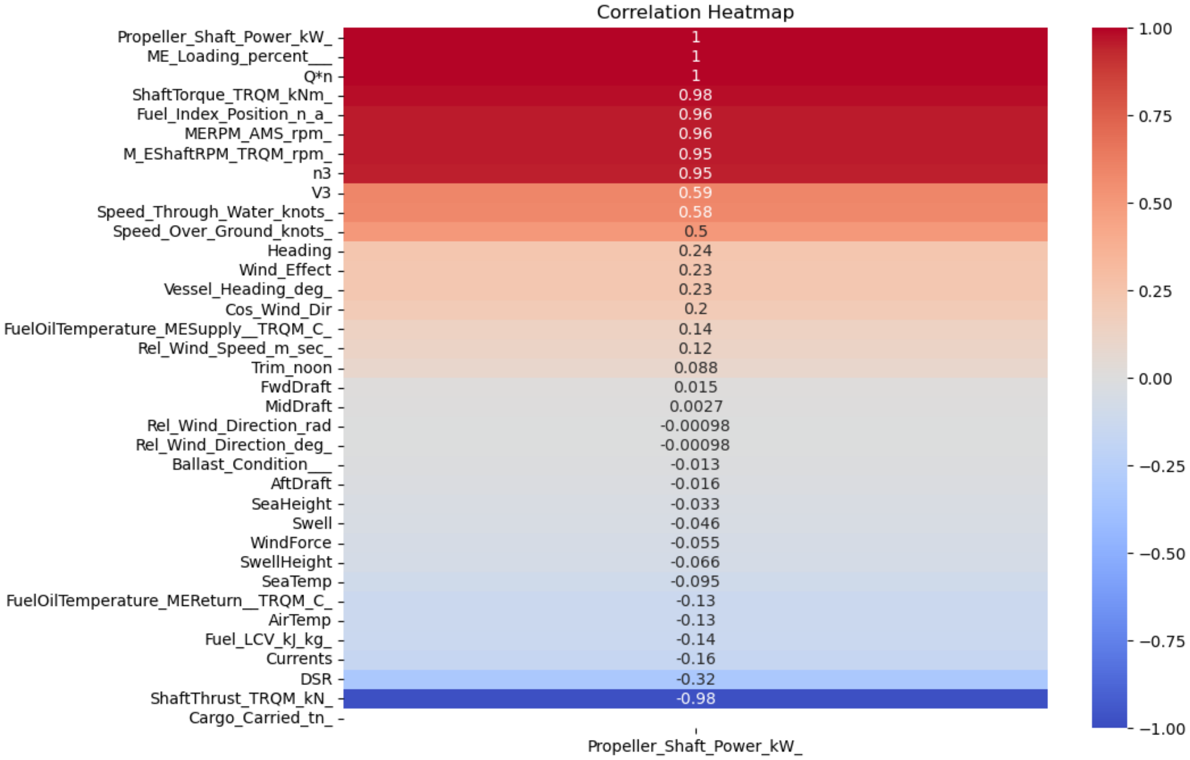


Figure 35: Pearson Correlation Coefficients values between input and output variables.

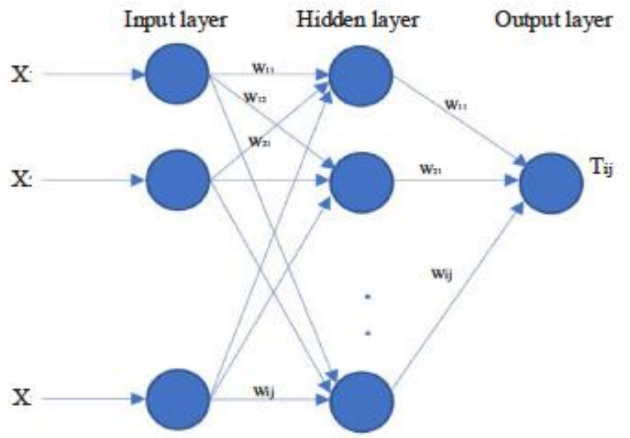
Hence, for the reasons described in **Feature Selection** paragraph, the following variables were excluded from the input variables selection process.

Highly Correlated Parameters
ME Loading Percentage (%)
Q*n
Shaft Torque (kNm)
Fuel Index Position
ME RPM (rpm)
n^3

Table 12: Highly correlated features, excluded from the input variables.

4.3.2 Random Forest regression classifier

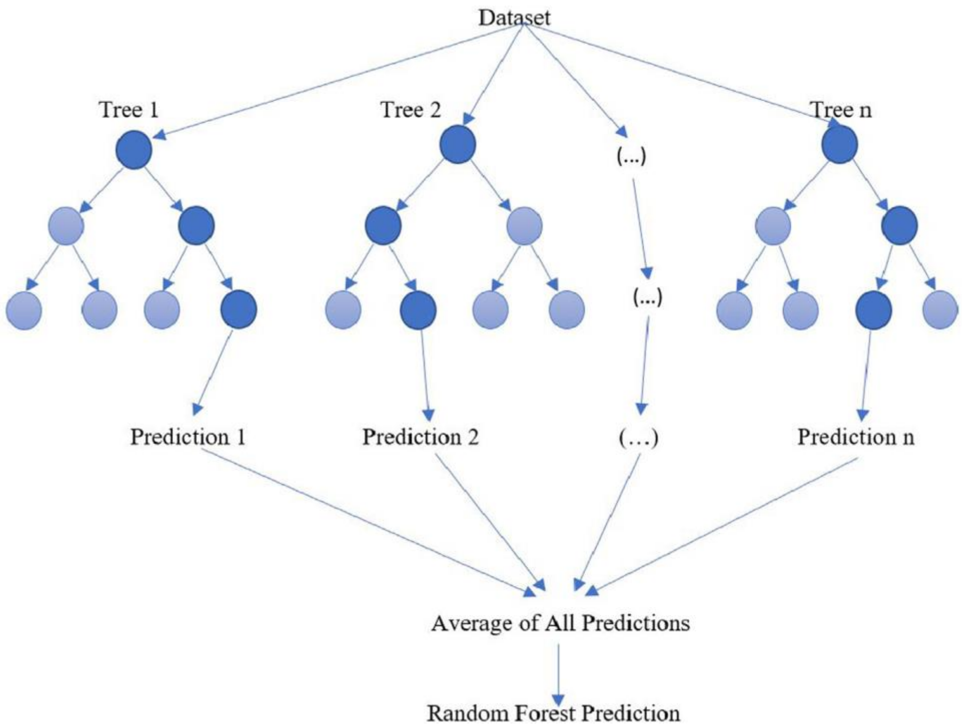
The Random Forest Regression is based on the decision tree regression method which is used on numerical data. Decision trees are constructed by multiple decision and leaf nodes according to the inputs and the outputs and are used in operations research, specifically in decision analysis, to help identify a strategy most likely to reach a goal.



(Uyanık et al., 2020)

Figure 36: Decision Tree Regression

To achieve higher accuracy and stability in prediction, the random forest algorithm generates numerous decision trees and aggregates them, as illustrated in Figure 37. Since the decision is based on multiple trees, the results are discrete. The bagging technique is employed to create the random forest for this regression model. In this technique, new trees are constructed by repeatedly sampling from the dataset, and the random forest is generated from these trees.



(Uyanık et al., 2020)

Figure 37: Random Forest Regression

Additionally, since Currents describe the difference between the SOG and STW features we will select two out of the tree variables, as goes the same for the Drafts and the Trim. The importance of each feature was validated through the Random Forest Regression model, giving the following results:

a/a	Parameters' Importance – Before repair	Parameter's Importance – After Repair
1	Wind Effect	Wind Effect
2	Mean Draft	Mean Draft
3	Trim	Trim
4	Sea Height	Sea Height
5	Speed Through Water (knots)	Speed Through Water (knots)
6	Currents	Currents
7	Heading	Heading
8	Swell Height	Swell Height
9	DSR	DSR
10	Fuel Oil Temp. (ME Supply)	Fuel Oil Temp. (ME Supply)
11	Fuel Oil Temp. (ME Return)	Fuel Oil Temp. (ME Return)

Table 13: Variable's importance enumeration for before & after repair datasets.

Additionally, through the Random Forest Regression Classifier, the Mean Squared Error (MSE) and R-squared values were calculated in regards to the number of inputs used to predict the PSP variable. To visualize the difference of these values compared with the number of inputs the following plots were created:

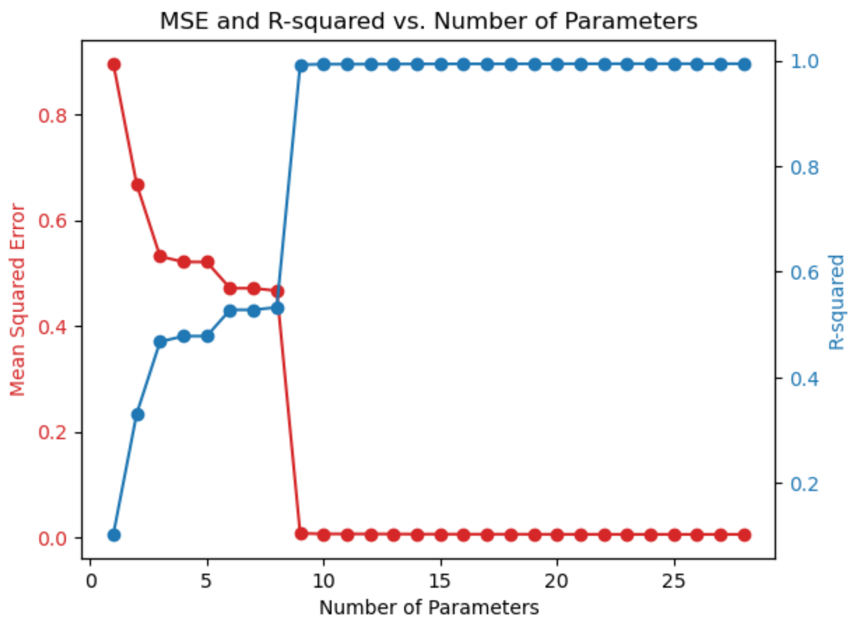


Figure 38: MSE & R-Squared values compared with the number of inputs used for the predictive model. (before repair)

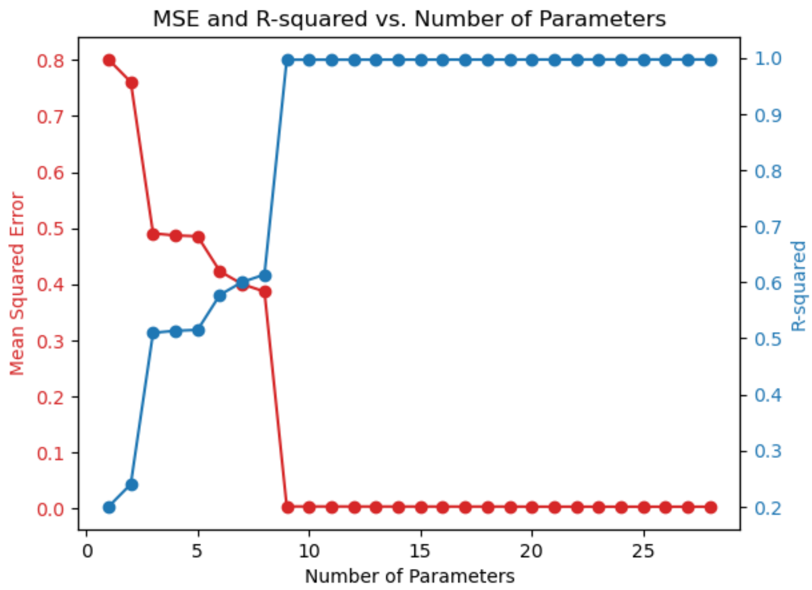


Figure 39:MSE & R-Squared values compared with the number of inputs used for the predictive model. (after repair)

It is clear that for both datasets the most efficient number of variables used as inputs is 9, since the MSE drops significantly but after that it remains stable and doesn't fluctuate. Hence, taking into account the importance calculated above (Table 13: Variable's importance enumeration for before & after repair datasets.), the final 9 variables used as inputs for the predictive models should be the following:

a/a	Parameters' Importance – Before repair	Parameter's Importance – After Repair
1	Wind Effect	Wind Effect
2	Mean Draft	Mean Draft
3	Trim	Trim
4	Sea Height	Sea Height
5	Speed Through Water (knots)	Speed Through Water (knots)
6	Currents	Currents
7	Heading	Heading
8	Swell Height	Swell Height
9	DSR	DSR

Table 14: Final parameters that will be used as inputs in the predictive models.

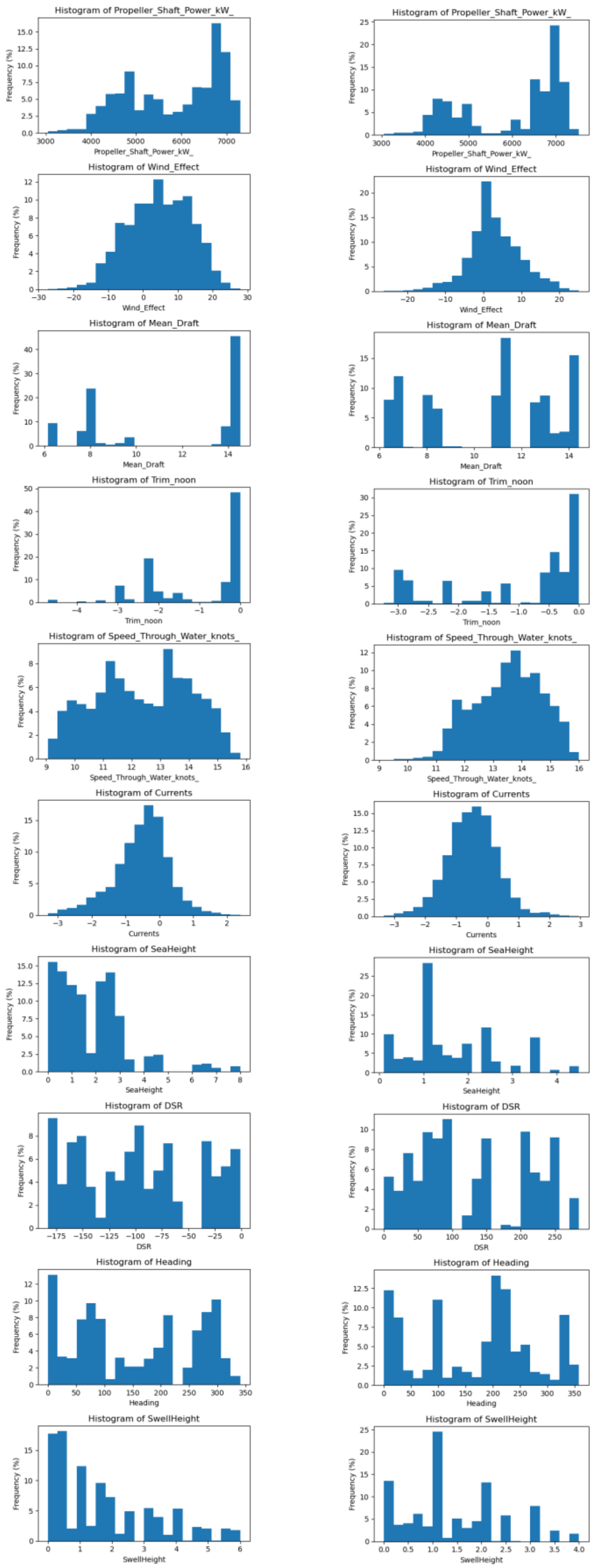


Figure 40: Model's input histograms (Final data) - before & after

5 Artificial Neural Networks

Deep learning methods are based on neural networks, also referred to as artificial neural networks (ANNs) or simulated neural networks (SNNs), which are a subset of machine learning. As Grossi & Buscema, (2008) mentioned, their generation was inspired by the structure and operation of the biological nervous system which consists of more than 10 billion neurons that interact with each other with more than 60 trillion connections. This structure provides the brain with the ability to process stimuli and shape rules through "experience". Hence as with the brain, all these layers of interconnected nodes (neurons) make up ANNs, which process information and are able to make decisions based on input data.

Artificial neural networks (ANNs) consist of interconnected nodes that form layers. A basic structure of an ANN consists of an input layer, one or more hidden layers, and an output layer. These nodes, or artificial neurons, are enabled through their own thresholds and associated weights. Specifically, if the output of a neuron is higher than the threshold value, it will activate and send data to the next layer of the network. Conversely, if the output is lower than the threshold value, no data will be transmitted.

Additionally, Grossi & Buscema, (2008) state that as the basic building block of ANNs is the artificial neuron, receiving multiple inputs but generating a single output. To activate the artificial neuron, each input is multiplied by its corresponding synaptic weight, and the sum of the products is then passed through the neuron's activation function to determine the output. As mentioned previously, ANNs contain many interconnected artificial neurons, which are organized into layers. One of the most important, or the most important, layer is the first layer also known as the input layer, due to the fact that is comprised by the selected input variables or features of the model. The second layer is the first hidden layer, whose neurons receive inputs from the first layer and output to the next layer. This process is repeated for each subsequent hidden layer until the last layer, whose output is the final output of the network. Although ANNs require training data to improve their accuracy, once the optimization of the learning algorithms is set, they can be powerful tools for classifying and clustering data at high speeds in computer science and artificial intelligence.

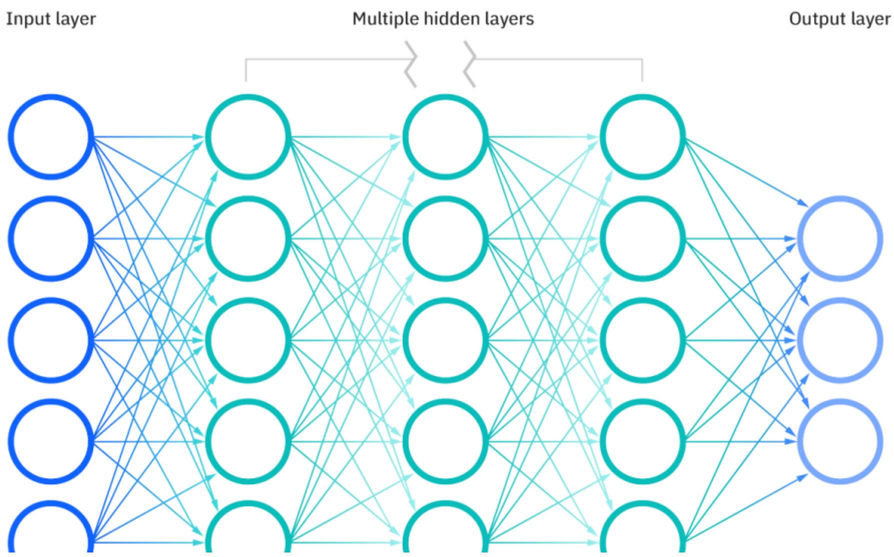


Figure 41: Basic Structure of ANNs (IBM, n.d.)

5.1 Basic hyperparameters of an ANN model

Hyperparameters are of major importance when designing an ANN model. The appropriate selection of the model's hyperparameters influence the performance of an artificial neural network (ANN), since they affect the settings that determine how the model learns from the input data. Moreover, the correct selection of hyperparameters could improve the model's accuracy, training convergence, and robustness as well as prevent the model's overfitting. Hence, it is important to select the best hyperparameters, to achieve the best possible performance. Some of the most important hyperparameters will be analyzed in the following paragraphs.

5.1.1 Input Layers

The initial layer of an artificial neural network (ANN) is called the input layer, and it serves to take in the input data. The dimensionality of the input data determines the number of neurons in the input layer. Each neuron in the input layer represents a specific feature of the input data, with the values of these neurons being determined by the corresponding features in the input data. For example, in the case of image data, each neuron in the input layer corresponds to a single pixel in the image. While the input layer doesn't perform any computations or manipulations on the input data, it plays a crucial role in transmitting the input data to the next layer of the network, which is typically a hidden layer. Finally, for the scope of this study, the input layers are set to be the features of the final data set, after the processing of the raw data.

5.1.2 Hidden Layers

The depth and complexity of an artificial neural network (ANN) model, as stated by (Dastres & Soori, 2021), is largely determined by the number of hidden layers, which is a hyperparameter. However, the optimal number of hidden layers can vary depending on the complexity of the input data and the problem being solved. For simple problems, a single hidden layer may suffice, while more complex problems may require multiple hidden layers. However, increasing the number of hidden layers does not always guarantee better performance, and in some cases, it can lead to overfitting.

The best number of hidden layers and neurons in each layer is typically determined through experimentation and fine-tuning. A common approach is to start with a small number of hidden layers and gradually increase their number until the desired level of performance is achieved. Balancing the complexity of the model with its generalization capabilities is essential to avoid overfitting and ensure that the model can perform well on new and unseen data.

5.1.3 Output Layer

The output layer is the last layer of an artificial neural network (ANN) model that produces the model's output. The number of neurons in this layer is determined by the nature of the problem being solved, such as binary classification, multi-class classification, or regression. To make the output values interpretable and usable, an activation function is typically applied to the output of the output layer. This function can vary depending on the type of problem being solved, and some commonly used activation functions include sigmoid, softmax, and linear functions. In summary, the output layer is a crucial component of an ANN model that produces the final output. The number of neurons in this layer is chosen based on the problem being solved, and an activation function is applied to normalize the output values.

5.1.4 Activation Functions

Activation functions play a vital role in artificial neural networks (ANNs) as they introduce non-linearity in the output of individual neurons or nodes. Non-linearity is essential for ANNs to model complex relationships between inputs and outputs. According to (Sharma et al., 2020), there are various activation functions that can be employed in ANNs, with some of the most widely used being the sigmoid, ReLU, tanh, and softmax functions.

The **sigmoid function** maps input values to a range between 0 and 1 and is typically utilized in binary classification problems to interpret the output as a probability. In artificial neural networks (ANNs), a neuron's activation is determined by the output of a sigmoid function, which can then be used to make decisions or passed on to the next layer of neurons. The sigmoid function's non-linear properties allow ANNs to model intricate relationships between input and output variables. However, the sigmoid function's output can become saturated, leading to insensitivity to input changes for very large or small inputs. This limitation can result in vanishing gradients, making it challenging to train deep neural networks. To address this issue, alternative activation functions, such as ReLU and its variations, have been developed.

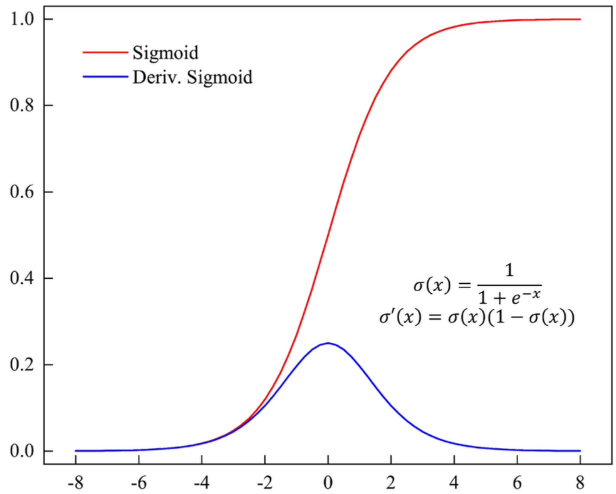


Figure 42: Simgoid activation function presentation. (Xiang et al., 2022)

The **Rectified Linear Unit (ReLU)** function is an activation function utilized in artificial neural networks (ANNs) to introduce non-linearity into the model. Its popularity stems from its simplicity, computational efficiency, and efficacy in deep neural networks. The ReLU function returns the input if it is positive, and 0 if it is negative. This enables faster and more efficient training of deep neural networks by reducing the vanishing gradient problem that can arise with other activation functions like the sigmoid function. The non-linear nature of the ReLU function allows ANNs to model complex relationships between input and output variables. However, the ReLU function is susceptible to the "dying ReLU" problem, which occurs when the input is negative, resulting in an output of 0 and a gradient of 0, thereby causing the neuron to no longer contribute to the learning process. This problem can be resolved by using variations of the ReLU function, such as the Leaky ReLU or the ELU, which have different approaches to handle negative inputs.

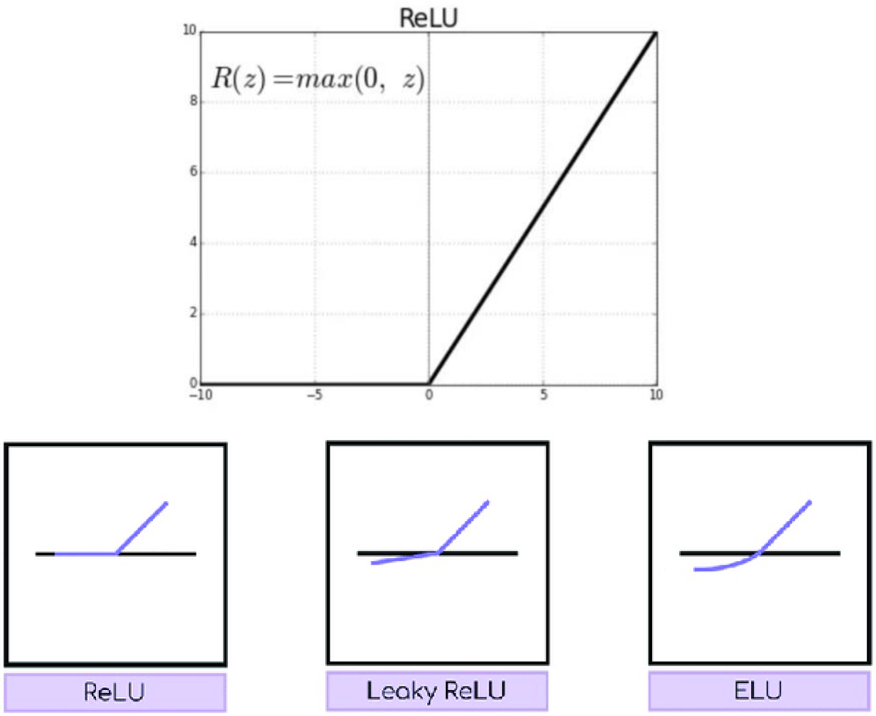


Figure 43: ReLU activation function variations presentations. (León et al., 2020)

The **tanh** function maps input values to a range between -1 and 1 and is similar to the sigmoid function but centered at 0. By using this activation function, negative inputs are strongly mapped as negative, and zero inputs are mapped close to zero, which is an advantage.

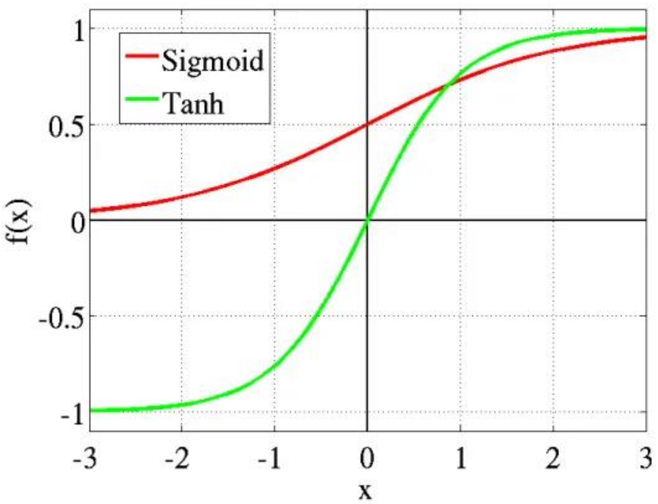


Figure 44: Sigmoid vs. tanh activation functions graphs. (Mukesh Chaudhary, 2020)

The **softmax** function is commonly used as the final activation function in ANNs that address multi-class classification problems as it maps each neuron's output to a probability distribution over possible output classes. The softmax formula is the following:

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_i}} \quad [5\text{-}1]$$

Where:

- $\vec{z}: The\ input\ vector\ to\ the\ softmax\ function$
- $z_i: The\ elements\ of\ the\ input\ vectors$
- $K: The\ number\ of\ classes\ in\ the\ multi-class\ classifier$

Other activation functions, such as the Gaussian, linear, and exponential functions, are less widely used but have some applications. The selection of an activation function depends on the specific problem and network architecture. In general, it is crucial to choose an activation function that is computationally efficient, produces desirable outcomes, and avoids issues like vanishing gradients or exploding gradients.

5.1.5 Loss Functions

The loss function in an Artificial Neural Network (ANN) model is responsible for evaluating the discrepancy between the predicted output and the actual output for a given input. The primary objective of the ANN is to minimize this loss function by modifying the network's weights and biases during the training process. The loss function, as stated by (Vishal Yathish, 2022) plays a crucial role in the training of the ANN model as it determines the performance and the rate at which the model can reach a solution. To be effective, a suitable loss function should be capable of detecting slight changes in the predicted output, easy to optimize, and appropriate for the particular problem being addressed. The most common Loss Functions are considered to be the following:

Mean Squared Error (MSE)

The **Mean Squared Error (MSE)** is a widely used loss function in an Artificial Neural Network (ANN) for regression tasks involving a continuous output variable. It calculates the mean of the squared difference between the predicted output and the actual output for a given input. The MSE formula is expressed as:

$$MSE = \left(\frac{1}{n}\right) * \Sigma(y_i - \hat{y}_i)^2 \qquad [5\text{-}2]$$

Where:

- $n: is\ the\ number\ of\ samples\ in\ the\ dataset.$
- $y_i: is\ the\ actual\ output\ for\ the\ i-th\ sample.$
- $\hat{y}_i: is\ the\ predicted\ output\ for\ the\ i-th\ sample.$

During the training phase of the ANN, the network's weights and biases are modified to minimize the MSE loss function. Lower MSE values indicate improved performance of the ANN as it signifies a closer match between the predicted and actual outputs. However, One drawback of this loss function is that it can be highly affected by outliers, meaning that data points with extreme values can have a disproportionate impact on the value of the loss function.

Root Mean Squared Error (RMSE)

The **RMSE** is a widely used loss function in Artificial Neural Networks (ANN) for regression tasks, similar to the MSE. However, the primary distinction between RMSE and MSE is that RMSE involves taking the square root of the mean of the squared differences between the predicted and actual outputs, resulting in an interpretable metric in the same units as the target variable. The RMSE formula is expressed as:

$$RMSE = \sqrt{\left(\left(\frac{1}{n}\right) * \Sigma(y_i - \hat{y}_i)^2\right)} \quad [5\text{-}3]$$

Where:
- n: $is\ the\ number\ of\ samples\ in\ the\ dataset.$
- y_i: $is\ the\ actual\ output\ for\ the\ i-th\ sample.$
- $\hat{y}_i$: $is\ the\ predicted\ output\ for\ the\ i-th\ sample.$

Mean Absolute Error (MAE)

Similar to MSE and RMSE, the Mean Absolute Error (MAE) is a loss function used to quantify the difference between the predicted output and the actual output for a given input in an Artificial Neural Network (ANN). However, unlike MSE and RMSE, which use the squared differences between the predicted and actual outputs, MAE uses the absolute difference. To calculate the MAE, we take the average of the absolute differences between the predicted and actual outputs for all samples in the dataset. The formula for MAE is as follows:

$$MAE = \left(\frac{1}{n}\right) * \Sigma|y_i - \hat{y}_i| \quad [5\text{-}4]$$

Lastly, the MAE is less sensitive to outliers than MSE, as it does not involve squaring the differences between the predicted and actual outputs. Additionally, it is more interpretable than MSE and RMSE since it is expressed in the same units as the target variable.

5.1.6 Model Regularization Methods

Regularization in Neural Networks is a set of techniques which are made to prevent the model from overfitting and improve the generalization of the model. Overfitting is a phenomenon that follows when the training procedure makes the model to learn the training data too well and lacks the ability to perform well on the unseen data. Regularization helps the model tackle this problem by setting additional constraints to the model during the training procedure. The most common regularization methods, according to (Nusrat & Jang, 2018) are described below:

L1 & L2 Regularization (L1/L2): These methods are also known as Weight Decay and they are entitled to adding penalty terms to the loss function of the model. This results into discouraging the generation of large weights in the network. Specifically, L1 regularization or Lasso Regularization sets the penalty to the sum of the absolute values of weights multiplied by the tuning parameter, which represents the regularization strength. Additionally, L2 regularization or Ridge Regularization sets the penalty equal to the sum of the squared values of weights multiplied by the tuning parameter.

Dropout Regularization: This method can randomly deactivate a merit of the input values during each training step. Hence, it prevents the network from depending entirely on a specific input variable and promotes learning from a all the features.

5.2 Model Design

Keeping in mind all the available options for hyperparameters explained in the previous section (5.1Basic hyperparameters of an ANN model) the next step is to decide which of them will be used for building the best possible ANN prediction model with the given dataset and inputs.

5.2.1 Data Normalization

In order to be able to make the most out of the final given data it is important to normalize the data. This procedure will ensure that all input features are equally evaluated. This will prevent the domination of input values on the model's learning process, and at the same time avoid bias towards features with larger values. Additionally, data normalization supports the stabilization of the model's learning procedure by avoiding large fluctuations in the model's weights and biases, known as backpropagation. Finally, the data normalization serves as a generalization of data which, in addition to the aforementioned advantages, the model will improve its performance giving better results. For the normalization of data, the min-max method was used, in every input.

$$X_{normalized} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad [5\text{-}5]$$

Where,

$$X_{\min} : The\ minimum\ value\ of\ X.$$
$$X_{\max} : The\ maximum\ value\ of\ X.$$

5.2.2 Data Shuffling

Shuffling the data is critical to prevent from results that came through biased train and test sets, which correspond to a specific time frame of the vessel's operation that do not represent the overall data distribution. Additionally, through shuffling we accomplish the elimination of patterns or trends, allowing a more efficient way of the model's generalization and learning procedures. In general data shuffling ensures that the model will learn and make predictions based on a more representative sample of data, enhancing its performance and accomplishing more reliable results.

5.2.3 Data Split

In order for a Neural Network to be able to provide predictions about its variables, the dataset must be divided into separate subsets for training, validation, and testing purposes. The most popular split is between a training set and a test set. For the purpose of this study, we divided the dataset into a subset for training and another for testing, following a split ratio of 64% training, 16% validation, and 20% testing.

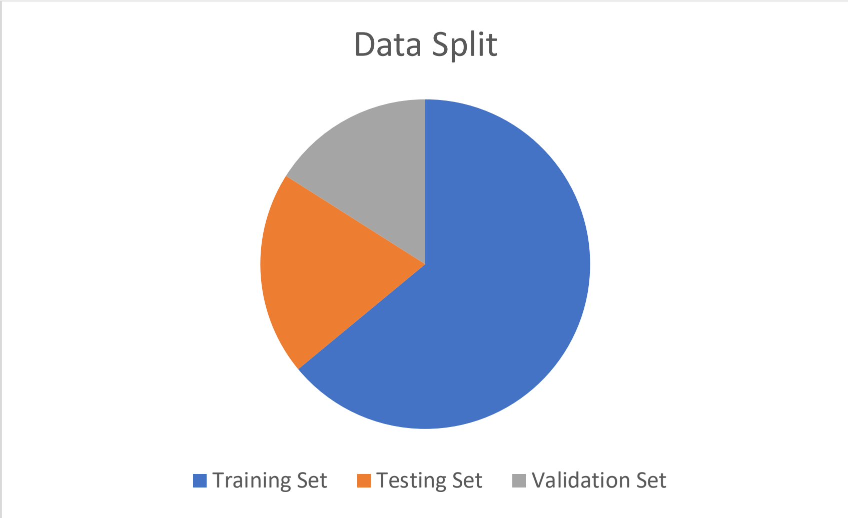


Figure 45: Data split between Training & Test sets.

5.2.4 Tuning the models hyperparameters

To decide which parameters should be used in the predictive model required a many tries and a lot of work. Each parameter was selected by experimenting with the model while changing the parameters and observing the impact that they had on the model's accuracy. That accuracy was calculated through the loss function which was set to be the MSE loss function (5.1.5 Loss Functions). Hence, the basic method followed was the following:

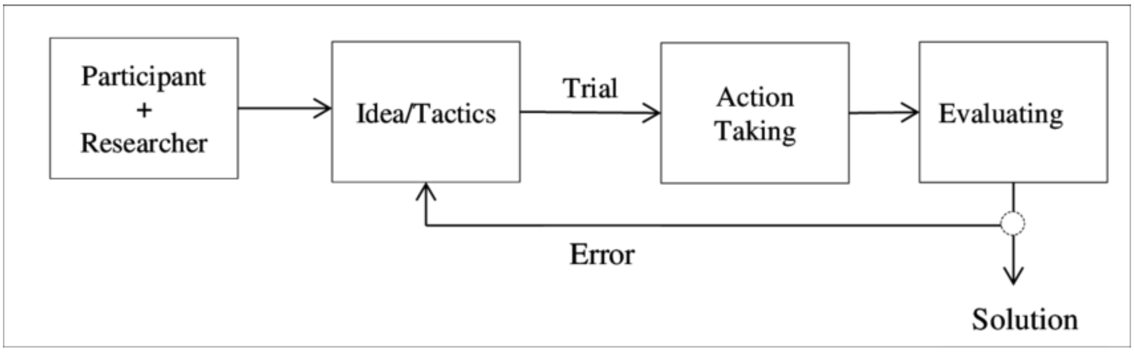


Figure 46: Trial-Error procedure diagram.

Being inspired by the Trial-Error procedure strategy many trials were executed to find the hyperparameters' consistency of the prediction model. Specifically, hidden layers were added to the model one by one, until the loss function (MSE) was seeing no improvement. Similarly, units were added to each layer until the error stabilized. The reason behind that strategy was to make sure that the predictive model would have the best accuracy possible while being as little as possible complicated.

Regarding the batch size and the number of epochs, they were found to be of major importance to the model's accuracy. The number of epochs defines the number of times that the training set will run through the model in order to get the information needed for the model to be best trained. Hence, increasing the number of epochs could eventually increase the model's accuracy, since it would be better trained. On the other hand, adding many epochs in the model could result in overfitting, where the model becomes too specialized to the training data and performs poorly to the unseen data.

As far as the batch size is concerned, it describes the number of samples that a predictive model processes before the model's parameters are updated through each training repetition and impacts the model's convergence and generalization. Having a larger batch size can provide more stable and smoother updates on the parameters of the model. Although, using larger batch sizes takes up more memory which may result in slower training processes, while also they could result in the model converging to a suboptimal solution.

Furthermore, regarding the activation functions, following the same plan as before all of the available functions, as mentioned in **5.1.4 Activation Functions**, were examined and the most promising one was selected.

To implement all these trials and to visualize their influence on the model's performance, a loss function plot was created, where each parameter was set to different values while the others remained the same. Hence, it was easier to understand how each hyperparameter performs, in terms of decreasing the loss function's values, as well as the duration of the training procedure. These Trial-Error procedures were applied on a Neural Network that receives the Dataset before the propeller repair as an input. Although, due to the similar nature of both Datasets, both Neural Networks will have the same hyperparameters which will be determined in the following paragraphs.

5.2.4.1 Before-Repair Model

Activation Function

In order to select the most suitable activation function for the model, it is important to look into the input datase**t** values. As presented in Figure 40: Model's input histograms (Final data) - before & after**,** the given Dataset which will function as the input values of the prediction model, doesn't contain negative values and this will be a decisive parameter for the selection of the activation function. As per **Activation Functions** paragraph, the most common activation function in machine learning is the ReLU function. Additionally, it was mentioned that it is prone to the ''Dying ReLU'' problem when the input dataset contains negative values. Although, since our Dataset has been normalized into positive values, this shouldn't influence our decision and hence we will use "ReLU" activation function. Finally, a basic model will be introduced in Table 15 which will be used as a basis to determine the other hyperparameters:

Basic Model	
Activation Function	ReLU
Hidden Layers	2
Units per Hidden Layer	64
Weights Initializer	Random
Optimizer	Adam
Epochs	50
Error Function	MSE

Table 15: Basic Model's Hyperparameters for the activation function selection.

Hidden Layers

According to the Trial-Error procedure, 7 models were created with each one of them having the same characteristics except for the number of hidden layers. Hence, through the learning curves of each model, we will be able to find the most suitable number of hidden layers for the ANN. The structure of the main model's hyperparameters is presented in the following table:

Basic Model	
Activation Function	ReLU
Units per Hidden Layer	64
Weights Initializer	Random
Optimizer	Adam
Epochs	50
Error Function	MSE

Table 16: Basic model's hyperparameters for the number of hidden layers selection.

Models	**No. of Hidden Layers**
Model 1	2
Model 2	3
Model 3	4
Model 4	5
Model 5	6
Model 6	7
Model 7	8

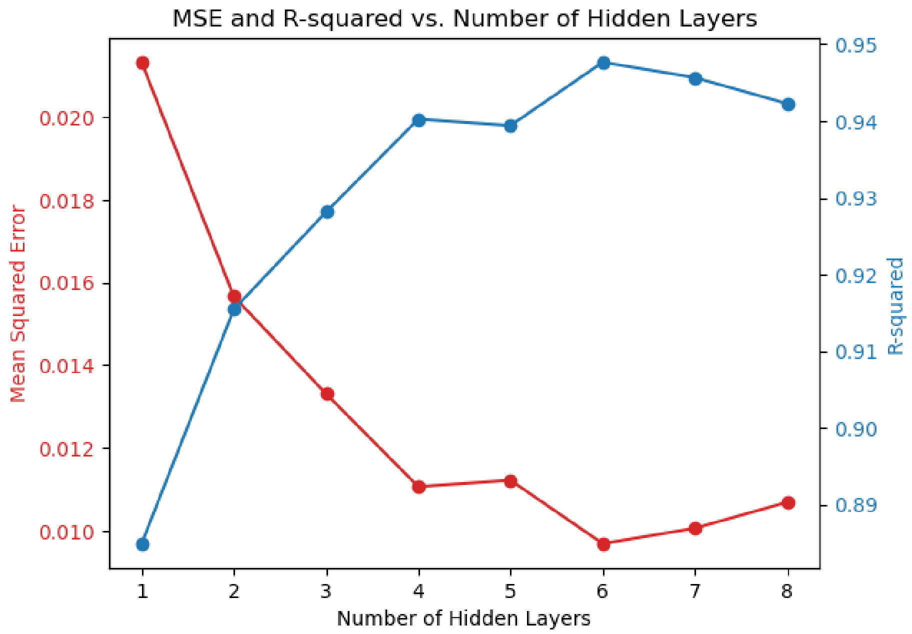


Figure 47: MSE & R2 values fluctuation in correspondence to the number of hidden layers of given model.

Through the observation of the MSE & R-squared curves shown in Figure 47, we can notice that by increasing the number of hidden layers we manage to decrease the MSE error function values. Although, after Model 5 the decrease of the MSE values are insignificant and hence in order to keep our ANN model as simple as possible it seems better to choose the Model's 5 number of hidden layers which is 6.

Number of Epochs & Batch size

Keeping the hyperparameters as selected above, there are several models created with different numbers of epochs and batch sizes to find the best combination for the final ANN prediction model. Hence, the following models were examined through their learning curves.

Basic Model	
Hidden Layers	6
Units per Hidden Layer	64
Activation Function	ReLU
Weights Initializer	Random
Optimizer	Adam
Error Function	MSE

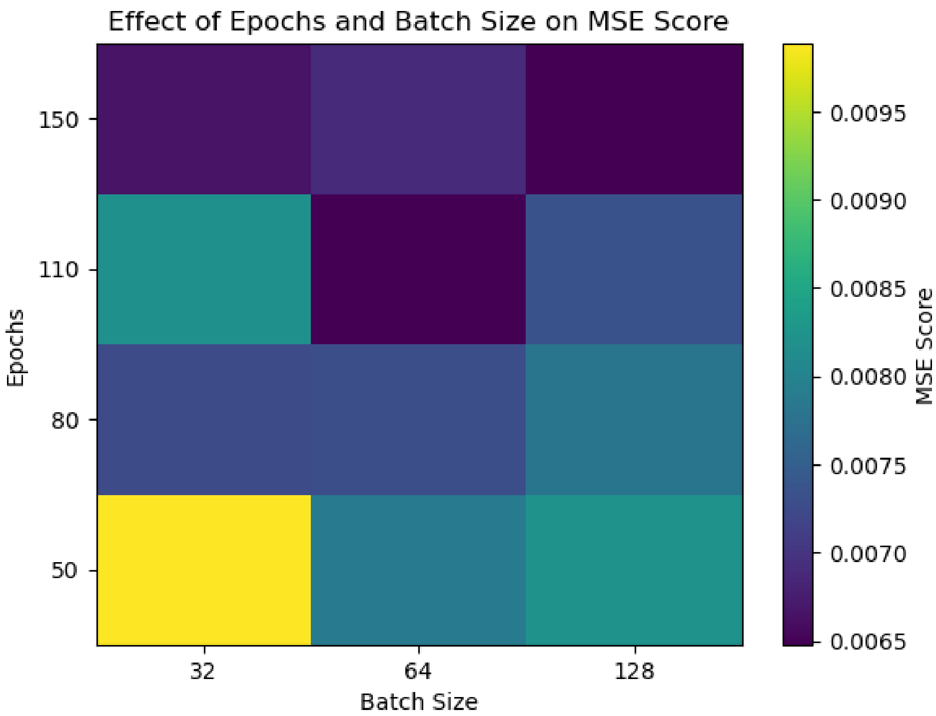


Figure 48: Effect of Epochs and Batch Size on models' MSE scores.

Based on the heatmaps above, the best-performing model seems to be the one with the 110 and 64 numbers of epochs and batch sizing respectively.

Number of units per hidden layer

In addition to the epochs and batch size numbers, the number of units inside each layer of the Network plays an important role in the model's performance and accuracy. To find the most suitable number of units per hidden layer, the Trial-Error procedure was followed for models whose number of units differed. Although, every model used the abovementioned hyperparameters as a basis. Hence, the following models were tested:

	Model 1	**Model 2**	**Model 3**	**Model 4**
Hidden Layers	6	6	6	6
Units per Hidden Layer	32	64	128	256

Epochs	110	110	110	110
Batch Size	64	64	64	64
Activation Function	ReLU	ReLU	ReLU	ReLU
Weights Initializer	Random	Random	Random	Random
Optimizer	Adam	Adam	Adam	Adam
Error Function	MSE	MSE	MSE	MSE

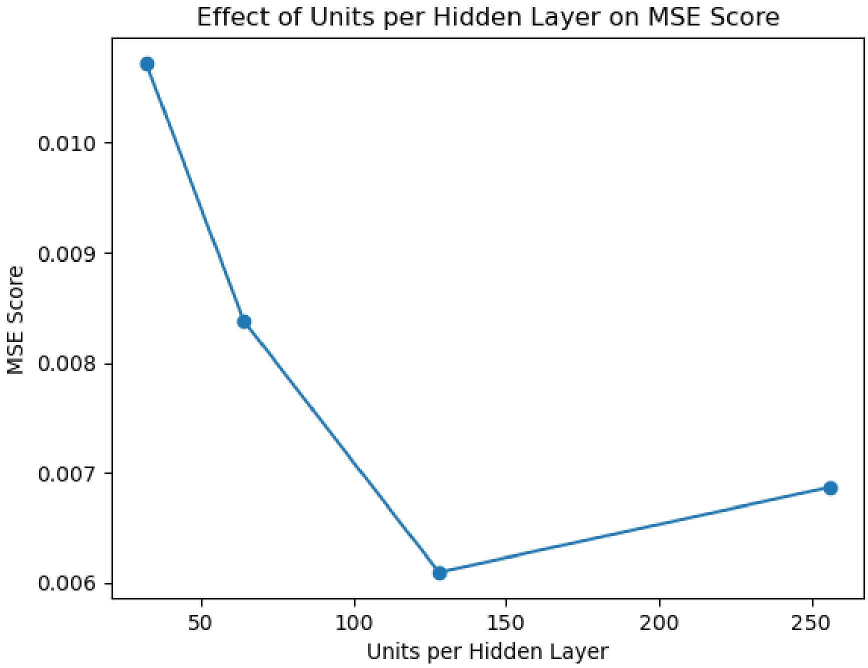


Figure 49: Effect of Units per Hidden Layer on models' MSE Score.

It is clear that based on Figure 49 the most efficient model is the one with 128 units per hidden layer. Although, after validating the model it was concluded that using that many units made the model to be unstable, hence the selected number was 64.

Model Normalization

To prevent the model from overfitting, selecting a method to normalize the weights and biases of the Neural Network is crucial. In order to select the most suitable method for the model, the Trial-Error procedure will be used once more. The following models are to be compared:

	Model 1	**Model 2**
Hidden Layers	6	6
Units per Hidden Layer	64	64
Epochs	110	110
Batch Size	64	64
Activation Function	ReLU	ReLU
Weights Initializer	Random	Random
Optimizer	Adam	Adam
Error Function	MSE	MSE
Regularization	L1 (0.001)	L2 (0.001)

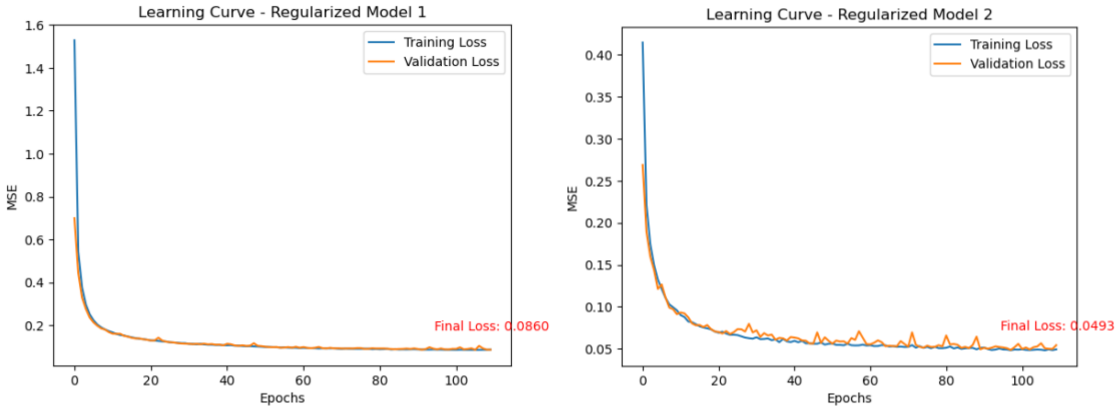


Figure 50: Learning curves of L1 & L2 regularized models.

Figure 50 illustrates the model's learning curves. The blue and orange curves represent the training and validation losses respectively during the training procedure of the model, calculated at the end of each epoch. Even though the L1 regularization method provides a more stable learning curve, the L2 method achieves better performance with a lower MSE score. Hence, since the model does not overfit, the L2 regularization method seems better for this study's model.

5.2.4.2 After-Repair Model

By following the same strategy of parameter evaluation, to the dataset after repair, the tuning of its hyperparameters was established. Hence, the final model used to predict the PSP feature of the dataset after the propeller repair is the following:

	Final Model – After Repair
Hidden Layers	5
Units per Hidden Layer	64
Epochs	110
Batch Size	64
Activation Function	ReLU
Weights Initializer	Random
Optimizer	Adam
Error Function	MSE
Regularization	L2 (0.001)

5.2.5 Model Training & Validation

5.2.5.1 Before-Repair Model

After the evaluation of each hyperparameter of the predictive model, the final model that will provide us with the wanted predicted values of the Propeller's Shaft Power is the following:

	Final Model – Before Repair
Hidden Layers	6
Units per Hidden Layer	64
Epochs	110
Batch Size	64
Activation Function	ReLU
Weights Initializer	Random
Optimizer	Adam
Error Function	MSE
Regularization	L2 (0.001)

Table 17: Final Model's hyperparameters.

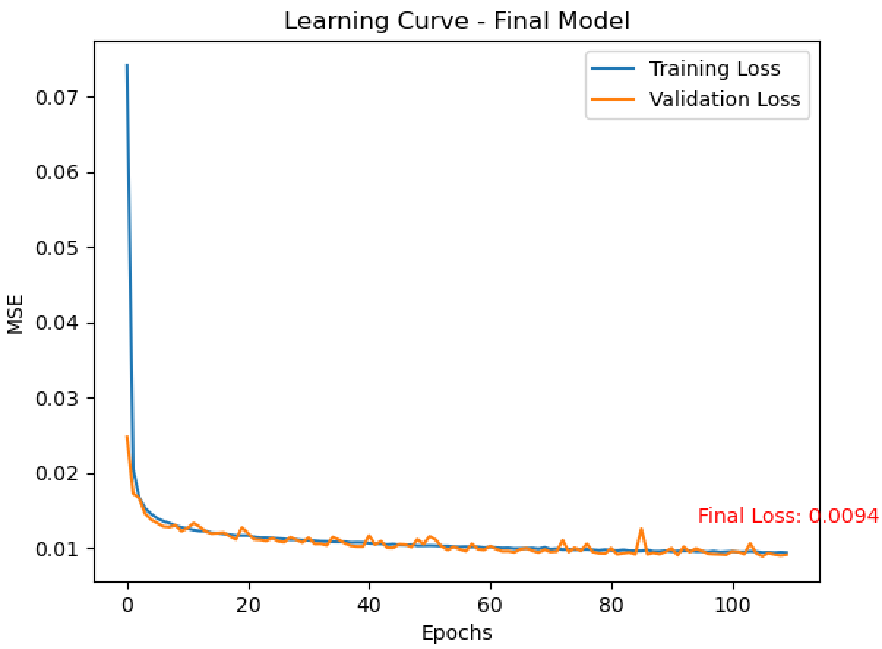


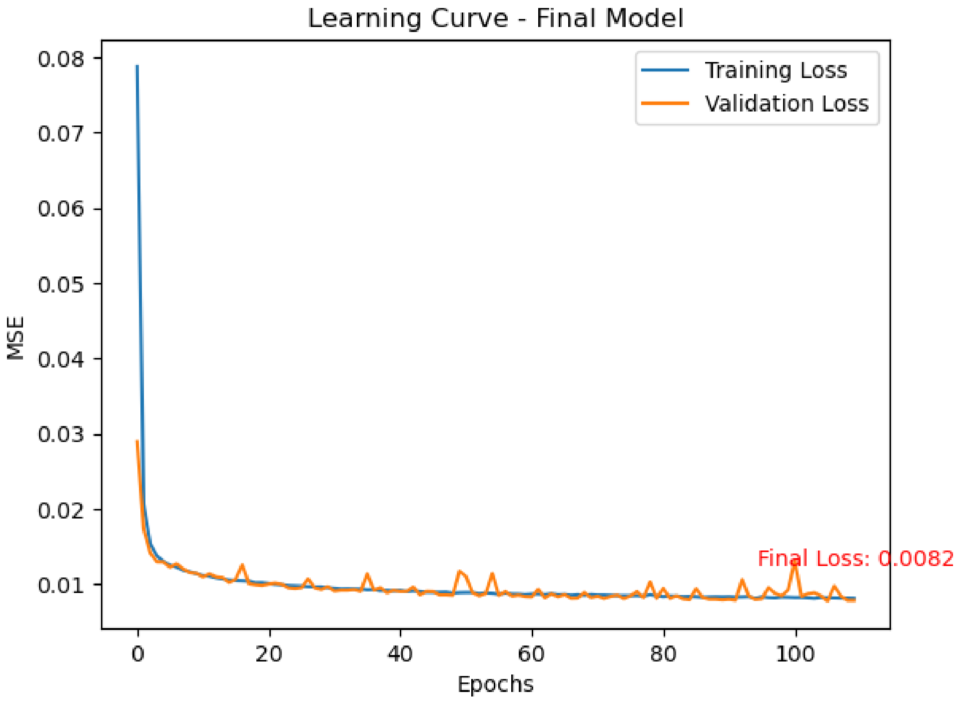
Figure 51: Final ANN model's learning curve.

As stated in Model Normalization chapter, and specifically in Figure 51, the final model's learning curve indicates a nearly linear behavior which decreases by the increase of epochs and eventually stops at MSE = 0.0094. Furthermore, the learning curve gives information about the model's fitting and shows that training loss is slightly less than the validation which proves that is a good fit. These features are a clear indication of the performance of the final ANN model which seems to be more than satisfactory.

5.2.5.2 After-Repair Model

	Final Model – After Repair
Hidden Layers	5
Units per Hidden Layer	64
Epochs	110
Batch Size	64
Activation Function	ReLU

Weights Initializer	Random
Optimizer	Adam
Error Function	MSE
Regularization	L2 (0.001)



Although the model seems a bit unstable, it manages to reach a low MSE value, equal to 0.0082, which will result in better model performance. The low stability of the model should be encountered by better data quality which would be possible if all continuous monitor sensors of the vessel worked properly.

5.2.6 Model Testing & Evaluation

After the tuning and the training of the model, using the Testing and Validation Datasets, it is important to test and evaluate the model's performance on a completely new and unknown dataset. This procedure will assess the trained model's performance and generalization capability along with its ability to handle unfamiliar data. Hence, this paragraph is determined to provide insights into the strengths and weaknesses of each model by presenting some statistical values along with the trajectories and the fitting of the model.

5.2.6.1 Dataset Before-Repair

	MSE (kW)^2	**RMSE (kW)**	**R^2**
Testing Dataset - Before	10152.4	100.8	0.979

Table 18: Statistical values of Model Before Repair.

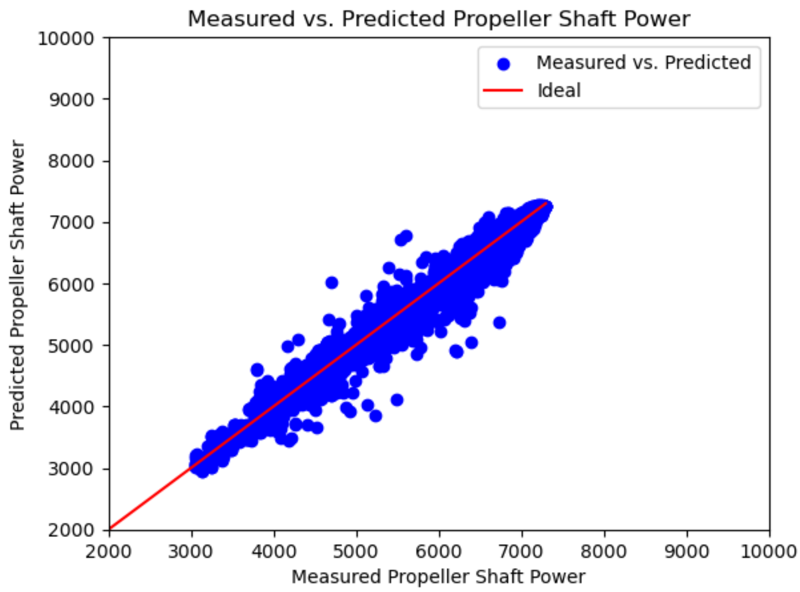


Figure 52: ANN Model Before-Repair: Measured vs. Predicted values.

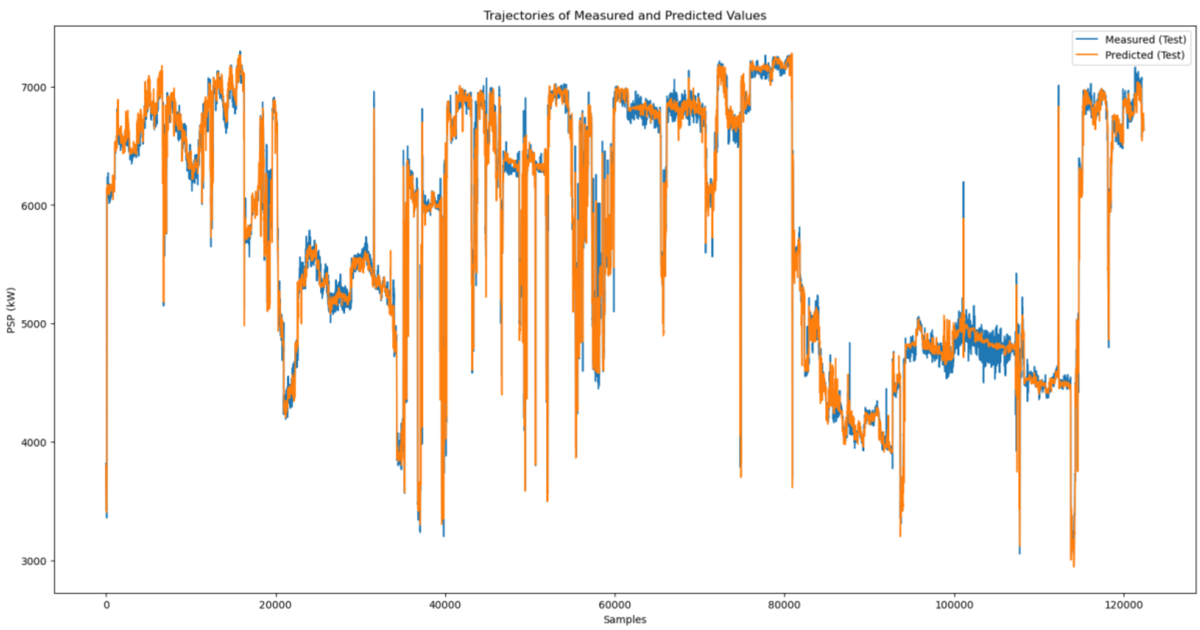


Figure 53: Trajectories of measured and predicted PSP values (Model Before Repair).

5.2.6.2 Dataset After-Repair

	MSE $(kW)^2$	RMSE (kW)	R^2
Testing Dataset - After	11125.8	105.5	0.989

Table 19: Statistical values of Model After Repair.

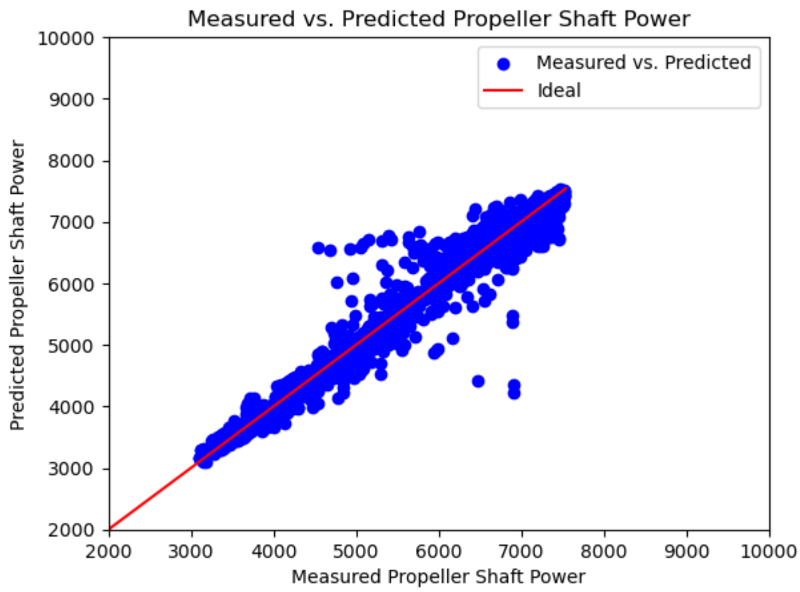


Figure 54: ANN Model After-Repair: Measured vs. Predicted values.

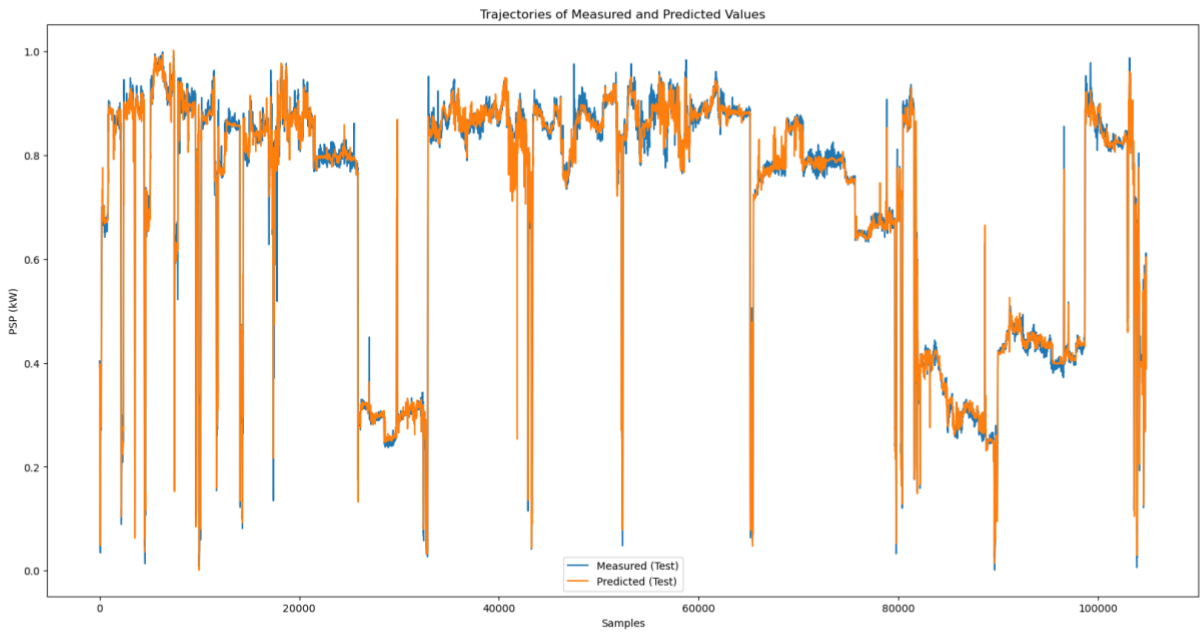


Figure 55: Trajectories of Measured and Predicted PSP values (Model After Repair).

Both models have reached the highest accuracy scores possible given the fact that the provided data did not have the best quality. However, from Figure 52Figure 53,Figure 54, andFigure 55 it is observed that in spite of 0.979 and 0.989 R^2 scores, both models react surprisingly well to unknown data, providing good performance.

6 Fouling Analysis

Marine fouling has become important to shipping companies since it affects the performance of the vessel significantly. As stated in 1.1.1 Vessel marine fouling there are several factors affecting it but one that does not depend neither on the vessel's characteristics or performance nor on the sea's and weather conditions is the number of days that have passed since the last cleaning. Hence, taking advantage of the ANN predictive models that were created in Chapter 5 in collaboration with the input feature of Days since Repair (DSR), it was possible to measure the impact of fouling in the performance deterioration of the vessel chronologically.

To do that, predictions based on synthetic test sets were made given in Table 20 & Table 21 with increasing DSR. All parameters were held constant, and they corresponded to the scantling draught condition which as per Figure 40: Model's input histograms (Final data) - before & after was the most common condition. Furthermore, predictions were made for a range of speeds through water between 9 to 16.5 knots to create power-speed curves which would assist us in comparing the performance of the vessel for different fouling conditions. To produce the power-speed curves, a cubic fit was applied to the predicted points since these curves would represent both the resulting predicted values trend as well as the empirical propeller law. The chosen model was the one created with data after the propeller repair which achieved better performance and accuracy scores. Such fouling analysis was applied for several weather and loading conditions and will be analyzed in the following subchapters.

6.1 Fouling analysis results

6.1.1 Scantling Loading Condition datasets

Prediction Set – Fouling Analysis	
Speed Through Water (knots)	$9 \leq V_s \leq 16.5$
Draft (m)	14.43
Trim (m)	-0.02
Currents	0
Sea Height	1.2
Wind Effect	2
Heading	200
Swell Height	1

Table 20: Synthetic dataset for fouling analysis prediction at scantling condition with calm weather.

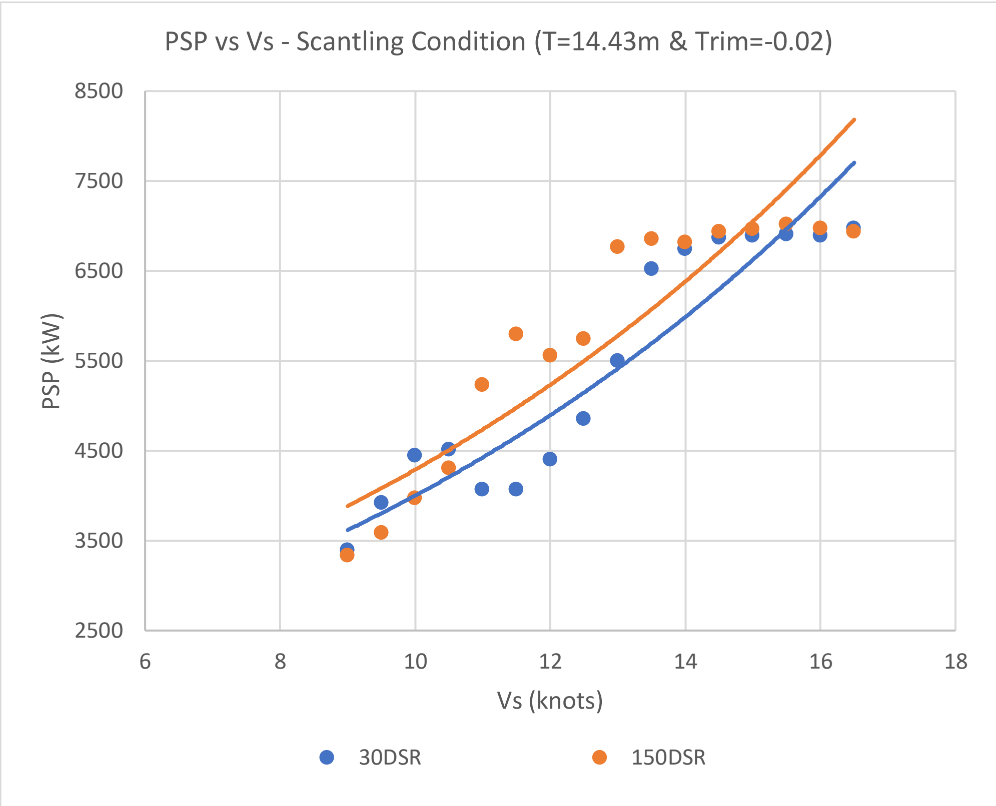


Figure 56: Speed-power curves for 30 and 150 DSR (Scantling condition with calm weather).

Prediction Set – Fouling Analysis	
Speed Through Water (knots)	$9 \leq V_s \leq 16.5$
Draft (m)	14.43
Trim (m)	-0.02
Currents	0
Sea Height	2.5
Wind Effect	10
Heading	200
Swell Height	1

Table 21: Synthetic dataset for fouling analysis prediction at scantling condition with rough weather.

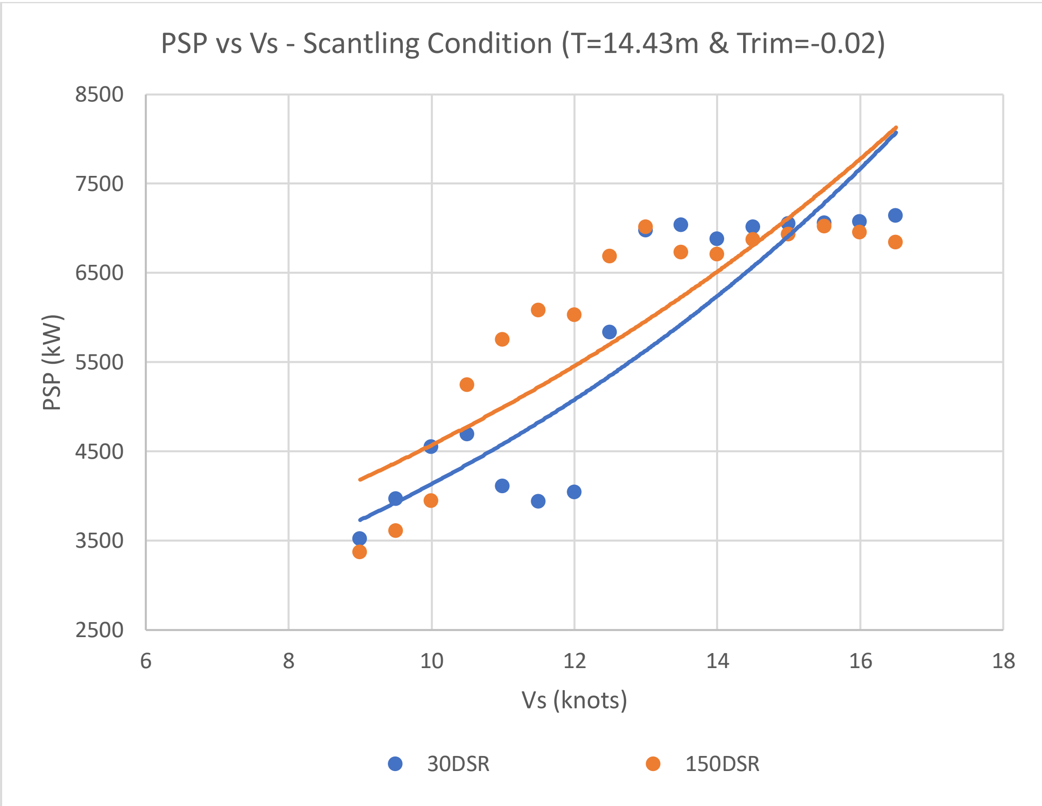


Figure 57: Speed-power curves for 30 and 150 DSR (Scantling condition with rough weather).

6.1.2 Design Loading Condition datasets

Prediction Set – Fouling Analysis	
Speed Through Water (knots)	$9 \leq V_s \leq 16.5$
Draft (m)	11.5
Trim (m)	-0.8
Currents	0
Sea Height	1.2
Wind Effect	2
Heading	200
Swell Height	1

Table 22: Synthetic dataset for fouling analysis prediction at design condition with calm weather.

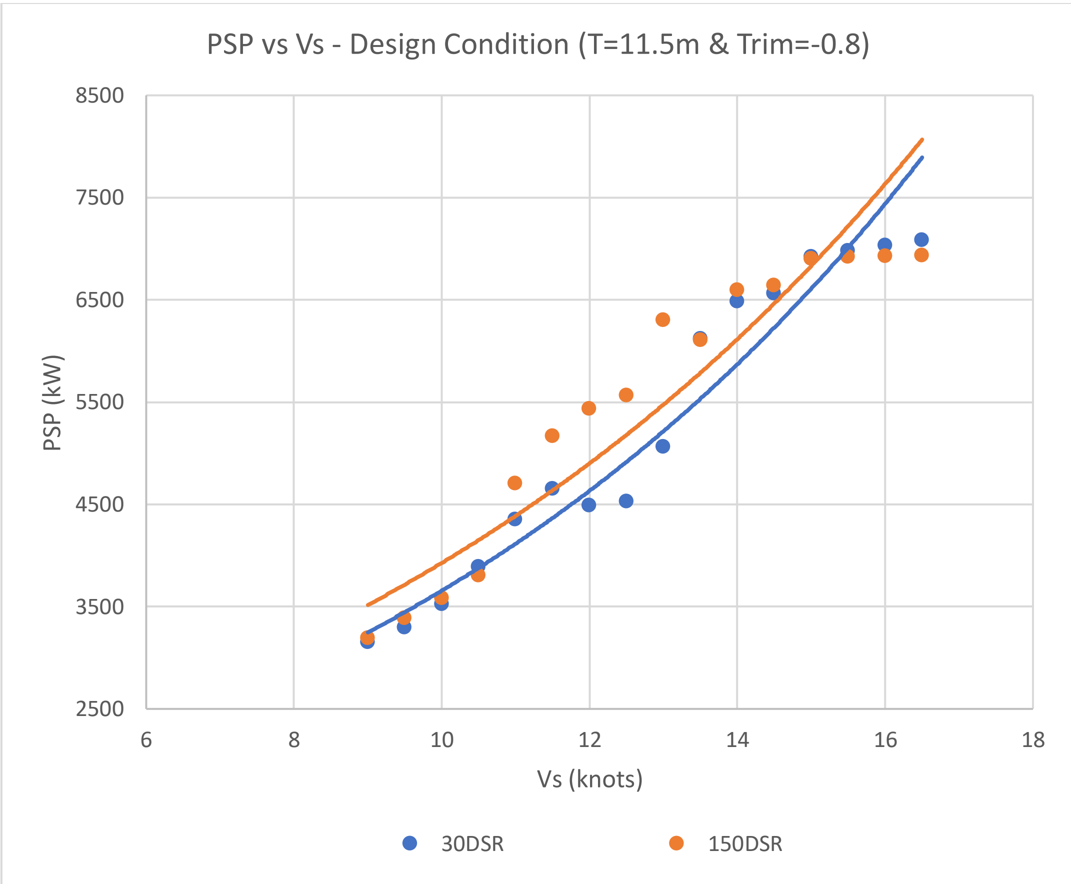


Figure 58: Speed-power curves for 30 and 150 DSR (Design condition with calm weather).

Prediction Set – Fouling Analysis	
Speed Through Water (knots)	$9 \leq V_s \leq 16.5$
Draft (m)	11.5
Trim (m)	-0.8
Currents	0
Sea Height	2.5
Wind Effect	10
Heading	200
Swell Height	1

Table 23: Synthetic dataset for fouling analysis prediction at design condition with rough weather.

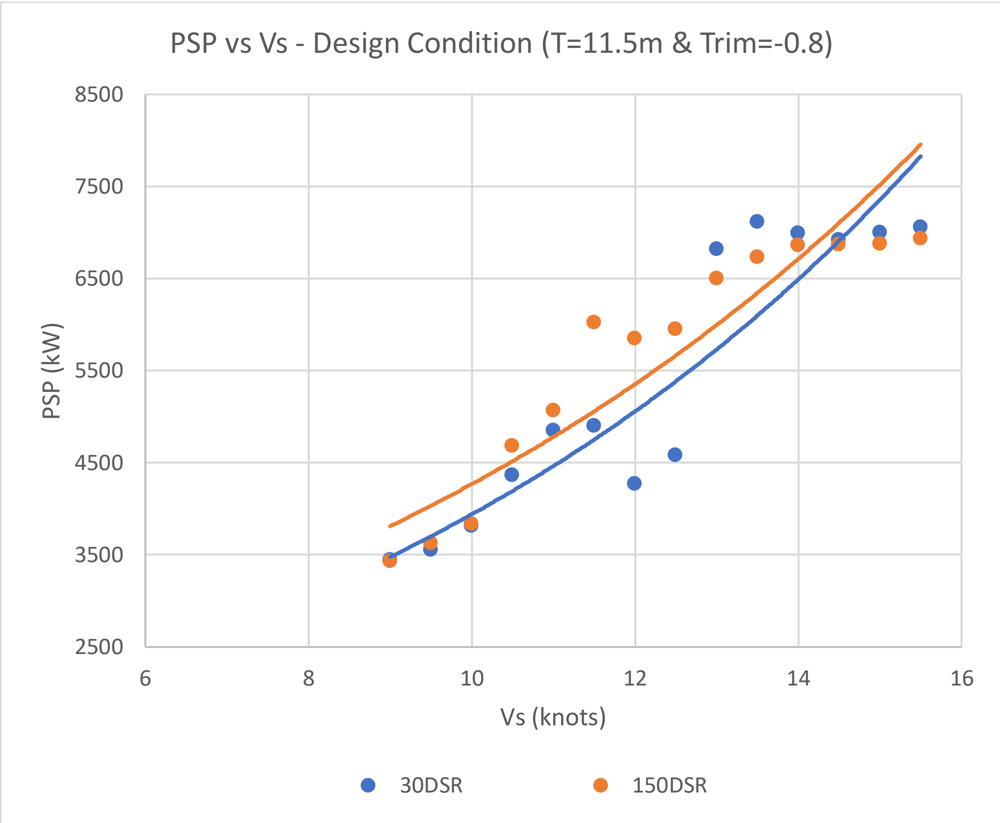


Figure 59: Speed-power curves for 30 and 150 DSR (Design condition with rough weather).

6.1.3 Ballast Loading Condition datasets

Prediction Set – Fouling Analysis	
Speed Through Water (knots)	$9 \leq V_s \leq 16.5$
Draft (m)	8
Trim (m)	-2.2
Currents	0
Sea Height	1.2
Wind Effect	2
Heading	200
Swell Height	1

Table 24: Synthetic dataset for fouling analysis prediction at ballast condition with calm weather.

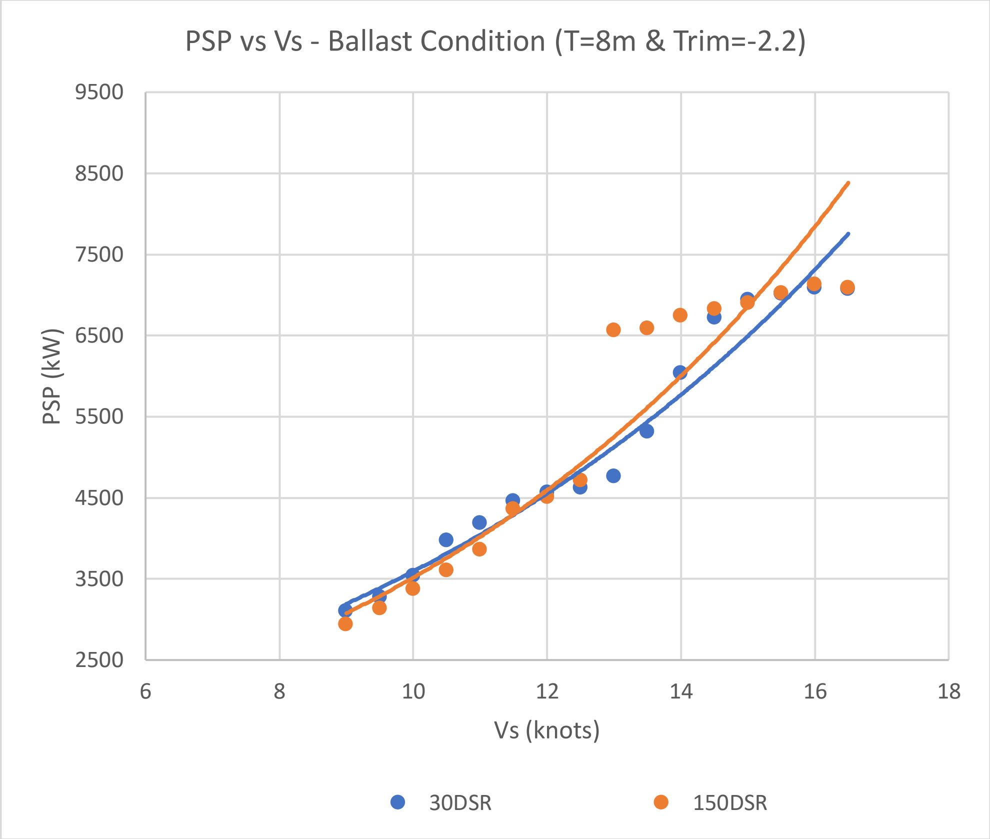


Figure 60: Speed-power curves for 30 and 150 DSR (Ballast condition with calm weather).

Prediction Set – Fouling Analysis	
Speed Through Water (knots)	$9 \leq V_s \leq 16.5$
Draft (m)	8
Trim (m)	-2.2
Currents	0
Sea Height	2.5
Wind Effect	10
Heading	200
Swell Height	1

Table 25: Synthetic dataset for fouling analysis prediction at ballast condition with rough weather.

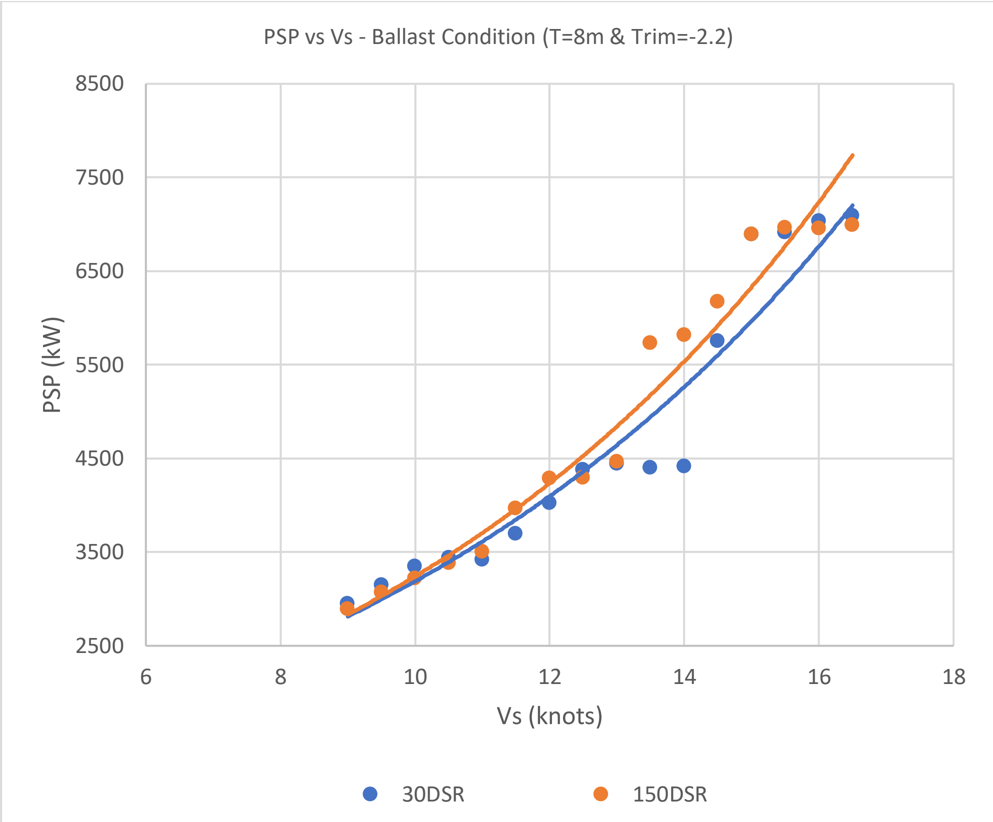


Figure 61: Speed-power curves for 30 and 150 DSR (Ballast condition with rough weather).

In every condition, it is noticed that the dataset which corresponds to 30 days since the repair has predicted lower PSP values for a specific range of speeds than the one corresponding to 150 days since the repair. This indicates the effect of marine biofouling on the vessel's hull and propeller's performance. In addition to that, according to the power-speed curves, the model seems to adjust correctly to the weather conditions differentiations since it predicts higher PSP values for rough weather than calm weather in every loading condition. To be more precise the following table represents the performance deviation between the power-speed curves of the two datasets in each condition.

Loading Condition	**Weather Condition**	**Mean PSP [kW]**	**Offset [kW]**	**Offset (%)**
Scantling	Calm	30DSR: 5439.55 150DSR: 5803.82	364.3	6.7 %
	Rough	30DSR: 5681.5 150DSR: 5988.5	307.5	5.4 %
Design	Calm	30DSR: 5258.81 150DSR: 5510.36	251.6	4.8 %
	Rough	30DSR: 5620.24 150DSR: 5814.72	194.5	3.5 %
Ballast	Calm	30DSR: 4711.22 150DSR: 4915.31	204.1	4.3 %
	Rough	30DSR: 5167.39 150DSR: 5335.25	167.9	3.2 %

Table 26: Performance deviation between 30DSR and 150DSR datasets for each loading and weather condition.

Calculating the performance deterioration in each condition lets the shipping companies evaluate the importance of the marine fouling effect on the vessel's performance and make decisions on whether a vessel needs to undertake cleaning procedures. Thereby, through the calculation of the extra power needed for the vessel to reach its working speed, shipping companies are able to monitor the additional fuel costs that correspond to that additional power. Hence, they are given the opportunity to forecast the most financially efficient time to clean their vessels which could eventually reduce their operational and fuel expenses and increase their profit.

In order to calculate the additional fuel costs (AFC) per hour that a shipping company would be entitled to, the following formula was applied:

$$\boldsymbol{AFC}(\$/\boldsymbol{day}) = \boldsymbol{F.C.} * \boldsymbol{SFOC} * \boldsymbol{Offset} \quad [6\text{-}1]$$

Where:

- **F.C. = 595($/ton)**: Average Fuel Cost at the moment.
- **SFOC (g/kWh)**: Fuel Oil Consumption corresponding to the mean ME Loading values of each condition presented in Table 26, based on the Main Engine's official shop tests results shown in Figure 62.
- **Offset (kW)**: Shaft power deviation between the 2 working conditions (30DSR vs 150DSR), presented in Table 26 ,and Figure 56, Figure 57, Figure 58, Figure 59, Figure 60, Figure 61 in regards to the loading condition.

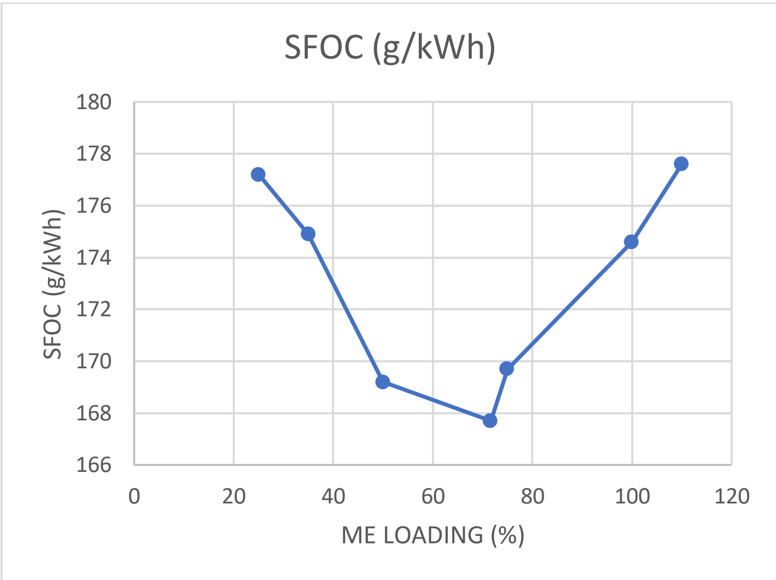


Figure 62: SFOC of vessel based on the ME Shop-Tests results.

Hence, the additional fuel cost for each weather and loading condition is presented below:

Loading Condition	**Weather Condition**	**ME Loading (%) [Mean STW]**	**SFOC [Mean STW]**	**Offset [kW]**	**AFC ($/day) [Mean STW]**
Scantling	Calm	55%	168.1	364.3	874 $/day
	Rough	60%	168.4	307.5	739 $/day
Design	Calm	53%	167.9	251.6	603 $/day
	Rough	58%	168.3	194.5	467 $/day
Ballast	Calm	48%	167.6	204.1	488 $/day
	Rough	53%	167.9	167.9	402 $/day

Table 27: Additional Fuel Cost between 30DSR and 150DSR datasets for each loading and weather condition on a mean STW.

Loading Condition	**Weather Condition**	**ME Loading (%) [12kn]**	**SFOC [12kn]**	**Offset [kW]**	**AFC ($/day) [12kn]**
Scantling	Calm	50%	167.7	372.3	891 $/day
	Rough	50%	167.7	457.5	1,094 $/day
Design	Calm	50%	167.7	302.6	723 $/day
	Rough	50%	167.7	232.5	556 $/day
Ballast	Calm	40%	167	102.1	243 $/day
	Rough	45%	167.4	35.9	84 $/day

Table 28: Additional Fuel Cost between 30DSR and 150DSR datasets for each loading and weather condition at STW=12knots.

Loading Condition	**Weather Condition**	**ME Loading (%) [15kn]**	**SFOC [15kn]**	**Offset [kW]**	**AFC($/day) [15kn]**
Scantling	Calm	70%	169.1	401.3	968 $/day
	Rough	70%	169.1	154.5	372 $/day
Design	Calm	69%	169.1	274.6	662 $/day
	Rough	70%	169.1	145.5	350 $/day
Ballast	Calm	69%	169.1	214.1	511 $/day
	Rough	70%	169.1	137.9	331 $/day

Table 29: Additional Fuel Cost between 30DSR and 150DSR datasets for each loading and weather condition at STW=15knots.

According to Table 27, Table 28, Table 29 it is noticed that the additional cost due to marine fouling development is something that shipping companies should consider since the amounts are significant. Hence, it would be wise for vessel operators to calculate these values to be able to find the most efficient period to conduct maintenance and hull cleaning operations and to minimize these additional fuel costs.

7 Conclusion & Future Suggestions

The objective of this study was to employ machine learning methods in order to build a predictive model about the vessel’s propeller shaft power and the affiliation of fouling and repairs with the vessel’s performance differentiation throughout the recorded period. This process was developed by utilizing a substantial dataset of operational data to highlight the importance of marine fouling both in the vessel’s performance but also the shipping company’s operational expenses. The given dataset was split into two subsets since the vessel had gone through a propeller repair during the recorded period. After a deep analysis of the data’s pre-processing procedures as well as the features’ engineering, a set of different neural networks were evaluated to find the best possible combination of hyperparameters that would ensure accurate outcomes. Throughout the pre-processing of the continuous monitoring system data, it was noticed that several sensors malfunctioned which led us to use a new dataset that consisted of both high and low-frequency data originating from the continuous monitoring systems and the noon reports respectively. This situation left us with no choice but to put all our effort to optimize both the pre-processing procedures, such as outlier detection, and the features engineering as well as the hyperparameters tuning of the model in order to be able to achieve the desired outcomes. For this reason, we developed methods to evaluate all the available methods of selecting both the input features and the model’s hyperparameters using a combination of a python-programmed Trial-Error method alternative, and Python’s built-in deep learning Random-Forest model. Thereby, despite the data anomalies that were presented the final models achieved high R^2 scores of 0.978 and 0.989 which is a clear indication of the pre-processing, features engineering, and hyperparameters optimization importance. The accuracy score difference between the two models can be explained by the difference in the amount of data that each model used to develop its predictive ability.

Finally, we put into use the model with the higher accuracy score and examined the effect of marine fouling on a vessel’s performance. This analysis was made by creating a time-dependent variable that connects the time elapsed since the propeller repair, named DSR (Days Since Repair). Furthermore, new datasets were created with different input values corresponding to different DSR, weather, and loading conditions and were applied to the selected model. The generated outputs indicated the conclusion that no matter the condition that a vessel sails, marine fouling will always increase its demanded propeller shaft power as time goes by. Hence, it is in the shipping company's best interests to monitor the vessel’s performance by using predictive models to calculate the most financially efficient period for its vessel to undergo maintenance and cleaning, preventing unexpected malfunctions and exorbitant amounts of fuel costs.

Future studies could use these conclusions and try to develop a model which would predict the most efficient timing for hull cleaning and maintenance operations considering the freight rates and the fuel cost as well as the average drydock or underwater hull and propeller cleaning costs at the time being. Thereby, shipping companies would be able to determine whether saving the additional fuel costs by performing premature maintenance actions would benefit them or not.

8 Bibliography

Adland, R., Cariou, P., Jia, H., & Wolff, F.-C. (2018). *The energy efficiency effects of periodic ship hull cleaning*. https://doi.org/10.1016/j.jclepro.2017.12.247

Anastasiou, T., & Thesis, D. (2022). *Development of Ship Performance models based on Artificial Neural Networks and Operational Data*. *March*.

Arndt, E., Robinson, A., Hester, S., Woodham, B., Wilkinson, P., Gorgula, S., & Brooks, B. (2021). *Factors that influence vessel biofouling and its prevention and management Final report for CEBRA Project 190803*.

Classification societies-what, why and how? (n.d.).

Dastres, R., & Soori, M. (2021). Artificial Neural Network Systems. *International Journal of Imaging and Robotics*, *21*, 13–25.

Digitalization in the maritime industry - DNV. (n.d.). Retrieved May 22, 2023, from https://www.dnv.com/maritime/insights/topics/digitalization-in-the-maritime-industry/data-usage.html

Farag, Y. B. A., & Ölçer, A. I. (2020). The development of a ship performance model in varying operating conditions based on ANN and regression techniques. *Ocean Engineering*, *198*, 106972. https://doi.org/10.1016/J.OCEANENG.2020.106972

Grossi, E., & Buscema, M. (2008). Introduction to artificial neural networks. *European Journal of Gastroenterology & Hepatology*, *19*, 1046–1054. https://doi.org/10.1097/MEG.0b013e3282f198a0

IBM. (n.d.). *What are Neural Networks? | IBM*. https://www.ibm.com/topics/neural-networks

Karagiannidis, P. (2019). *Data-driven Ship Propulsion modeling with applications in the Performance Analysis and Fuel Consumption prediction*. *July*. Diploma Thesis, NTUA.

Karagiannidis, P., & Themelis, N. (2021). *Data-driven modelling of ship propulsion and the effect of data pre-processing on the prediction of ship fuel consumption and speed loss*. https://doi.org/10.1016/j.oceaneng.2021.108616

Kim, Y. R., Jung, M., & Park, J. B. (2021). Development of a fuel consumption prediction model based on machine learning using ship in-service data. *Journal of Marine Science and Engineering*, *9*(2), 1–25. https://doi.org/10.3390/jmse9020137

Lang, X., Wu, D., & Mao, W. (2022). Comparison of supervised machine learning methods to predict ship propulsion power at sea. *Ocean Engineering*, *245*(December 2021). https://doi.org/10.1016/j.oceaneng.2021.110387

Laurie, A., Anderlini, E., Dietz, J., & Thomas, G. (2021). Machine learning for shaft power prediction and analysis of fouling related performance deterioration. *Ocean Engineering*, *234*(February), 108886. https://doi.org/10.1016/j.oceaneng.2021.108886

León, J., Escobar, J. J., Ortiz, A., Ortega, J., González, J., Martín-Smith, P., Gan, J. Q., & Damas, M. (2020). Deep learning for EEG-based motor imagery classification: Accuracy-cost trade-off. *PLoS ONE*, *15*(6). https://doi.org/10.1371/JOURNAL.PONE.0234178

Mukesh Chaudhary. (2020, August 28). *Activation Functions: Sigmoid, Tanh, ReLU, Leaky ReLU, Softmax | by Mukesh Chaudhary | Medium*. https://medium.com/@cmukesh8688/activation-functions-sigmoid-tanh-relu-leaky-relu-softmax-50d3778dcea5

Nusrat, I., & Jang, S. B. (2018). A comparison of regularization techniques in deep neural networks. *Symmetry*, *10*(11). https://doi.org/10.3390/sym10110648

Rochim, A. F. (2016). *Chauvenet's and Peirce's Criterion (literature review) IT-Governance View project IT Governance View project*. https://www.researchgate.net/publication/299829851

Sharma, S., Sharma, S., & Athaiya, A. (2020). ACTIVATION FUNCTIONS IN NEURAL NETWORKS. *International Journal of Engineering Applied Sciences and Technology*, *4*, 310–316. http://www.ijeast.com

Singh, K., & Upadhyaya, S. (2012). *Outlier Detection: Applications And Techniques*. www.IJCSI.org

Skamagkas, A. (2022). *UNIVERSITY OF PIRAEUS DEPARTMENT OF MARITIME STUDIES MSc in SHIPPING MANAGEMENT "AN APPLICATION OF DELPHI METHOD ON VESSEL PERFORMANCE MONITORING & OPTIMISATION."*

Statistical Rejection of "Bad" Data-Chauvenet's Criterion. (n.d.).

Uyanık, T., Karatuğ, Ç., & Arslanoğlu, Y. (2020). Machine learning approach to ship fuel consumption: A case of container vessel. *Transportation Research Part D: Transport and Environment*, *84*(May). https://doi.org/10.1016/j.trd.2020.102389

Valchev, I., Coraddu, A., Kalikatzarakis, M., Geertsma, R., & Oneto, L. (2022). Numerical methods for monitoring and evaluating the biofouling state and effects on vessels' hull and propeller performance: A review. In *Ocean Engineering* (Vol. 251). Elsevier Ltd. https://doi.org/10.1016/j.oceaneng.2022.110883

Vishal Yathish. (2022). *Loss Functions and Their Use In Neural Networks | by Vishal Yathish | Towards Data Science*. https://towardsdatascience.com/loss-functions-and-their-use-in-neural-networks-a470e703f1e9

Xiang, C., Wang, D., Pan, Y., Chen, A., Zhou, X., & Zhang, Y. (2022). Accelerated topology optimization design of 3D structures based on deep learning. *Structural and Multidisciplinary Optimization*, *65*(3). https://doi.org/10.1007/s00158-022-03194-0

Yusim, A. K., & Utama, I. K. A. P. (2017). An Investigation Into The Drag Increase on Roughen Surface due to Marine Fouling Growth. In *The Journal for Technology and Science* (Vol. 28, Issue 3).